UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE QUÍMICA E BIOQUÍMICA

Investigations of S100 proteins and their role in huntingtin aggregation

Marta Sofia Silva da Cruz Botelho

Mestrado em Bioquímica e Biomedicina

Dissertação orientada por:
Professor Doutor Cláudio M. Gomes
Professor Doutor Erich E. Wanker

2023

Acknowledgments

First and foremost, I would like to express my deepest gratitude to Prof. Dr. Cláudio M. Gomes, for giving me the opportunity and privilege to carry my Master's Thesis within his laboratory and research group, for his incredible and contagious scientific spirit and for all support and mentorship throughout this amazing year. I am also very thankful for the opportunity of developing part of my thesis in the MDC Institute in Berlin, under the supervision of Prof. Dr. Erich Wanker, to whom I am very thankful, for the warm welcoming, for accepting me in an amazing group and whose mentorship has not only enriched my academic knowledge but has also fostered personal and professional growth.

I would like to thank all Proteomics and Molecular Mechanisms of Neurodegenerative Diseases lab, for welcoming me like one of your own, and for giving me the space to be myself and allow me to grow in a different place outside home, specially to Roxanne, Megan, Orchid, Emily, Eduardo, Anabelle and Leo. I also want to thank to Anne and Martina for guiding me during this time.

I am very grateful for Prof. Dr. Bárbara Henriques and all my colleagues at the Protein Misfolding and Amyloids in Biomedicine laboratory. Thank you for your patience, discussions, input, guidance, and your friendship throughout my Master's thesis. All the conversations, laughs, difficulties and companionship shared were precious, and I feel extremely grateful that I was lucky enough to meet such wonderful and inspiring colleagues like you, Joana, Margarida, Mariana, Daniela, António and Guilherme. I will miss you all.

I would like to extend my thanks to Prof. Dr. Federico Herrera from the Cell Structure and Dynamics laboratory, for all the insightful discussions, for all the incredible support and for all the learning experiences. To all my fellow colleagues from the CSD, I would also like to express my thanks for receiving me so well in your laboratory, for the help and for the amazing conversations shared throughout my stay there. Thank you Ana Sofia, Brenda, Luana, André, Pedro and Fernanda.

A very special thanks to my friends, Maria Ramón, for the unconditional friendship shared for the past years, and for making my academic path an incredible experience, Mariana Barata, for being my oldest and best friend, and for the never-ending encouragement for all the little steps in my life, and to Maria Machado, for the unwavering support, inspiration and the cherished friendship we've been sharing throughout the years.

I'm at a loss for words for how grateful I feel for having such supportive parents, that have always given me the tools for becoming everything I wanted in life and for being my guiding light. For believing and loving me unconditionally. I want to take a moment to express my heartfelt gratitude and appreciation for my little sister, Mariana. Throughout the ups and downs of life, you have been a constant source of love, support, and joy.

Last, but certainly not least, I would like to thank Gonçalo, for almost a decade of love and companionship, that I believe can only be found once in a lifetime.

The research for this thesis was conducted at the Protein Misfolding and Amyloids in Biomedicine laboratory, affiliated with the BioISI - Biosystems & Integrative Sciences Institute within the Department of Chemistry and Biochemistry at the Faculdade de Ciências, Universidade de Lisboa (Ciências/ULisboa). This work was carried out under the guidance of Professor Cláudio Gomes. Additionally, a portion of the research was conducted at the Proteomics and Molecular Mechanisms of Neurodegenerative Diseases laboratory, affiliated with the MDC – Max Delbrück Center for Molecular Medicine, under the supervision of Professor Erich Wanker, during an Erasmus+ internship lasting 3 months. Financial support for this research was provided by FCT-Portugal (BD/11023/2022 MCS) and UID/MULTI/04046/2020 (BioISI), in the form of a BII – BioISI Junior Program fellowship awarded to M.B. The project also received funding from TWIN2PISPA (Twinning for Excellence in Biophysics of Protein Interactions & Self-Assembly), with grant GA101079147, under the EU-TWIN2PIPSA initiative, hosted by FC.ID/ULisboa. More information about TWIN2PISPA can be found at https://twin2pipsa.campus.ciencias.ulisboa.pt/

Abstract

Huntington's disease (HD) is a progressive neurodegenerative disorder characterized by an autosomal dominant inheritance, leading to a triad of motor, cognitive, and behavioural symptoms. HD is an incurable disease caused by a CAG trinucleotide repeat expansion in the *HTT* gene, which translates into an expanded polyglutamine tract (polyQ) in the huntingtin (Htt) protein, leading to the formation of mutant huntingtin (mHtt). The expanded polyQ tract destabilizes the conformation of huntingtin, and promotes the assembly of amyloidogenic protein aggregates, primarily composed of N-terminal Htt fragments. This accumulation disrupts various essential brain functions, ultimately leading to neuronal degeneration. S100B is a calcium-binding alarmin that is predominantly synthesized by astrocytes in the nervous system. Depending on the concentration, S100B operates as an aggravating pro-inflammatory cytokine or as a protective anti-aggregation chaperone, inhibiting amyloid aggregation and mitigating toxicity in the presence of Ca^{2+}. Considering S100B's significance in other neurodegenerative conditions and its chaperone role, this work aims to explore its potential involvement in huntingtin aggregation and in HD pathology. To address this, both cellular assays and *in vitro* assays were performed. First, interactions between S100B and huntingtin were investigated using a dual luminescence-based co-immunoprecipitation assay (LuTHy), for detection of protein-protein interactions in mammalian cells. Secondly, the effects of endogenous and exogenous S100B on live cells expressing wild-type and mutant huntingtin were studied through bimolecular fluorescence complementation (BiFC) assays, to evaluate potential changes in the amounts and dynamics of Htt inclusions. Furthermore, *in vitro* assays explored how the S100B protein influences Htt aggregation, employing a FRET-based Htt aggregate seeding (FRASE) approach, which revealed that S100B significantly delays and inhibits Htt aggregation, while operating at suprastoichiometric concentrations. With this work, it is expected to unravel the relationship between HD and the S100B protein, which might constitute a foundation for the development of novel therapeutic approaches.

Keywords: Huntington’s Disease, Huntingtin, protein aggregation, S100B.

Resumo

O sistema nervoso humano, é um atributo extraordinário e distintivo da nossa existência, destacando a singularidade da nossa espécie no mundo natural. Trata-se de uma rede altamente complexa composta por neurónios e células de suporte, contendo aproximadamente 86 mil milhões de neurónios que intercomunicam através de um vasto número de sinais eletroquímicos, em resposta a estímulos ou *inputs* sensoriais. Durante o desenvolvimento, os neurónios sofrem um processo designado de diferenciação terminal, tornando-se células pós-mitóticas que perdem a capacidade de se dividir e proliferar. Assim, estas células devem subsistir ao longo da vida de um indivíduo. A inevitável morte celular, associada à idade, resulta por si só na natural diminuição do número de neurónios no cérebro. Doenças neurodegenerativas, como a doença de Alzheimer (DA), a doença de Parkinson (DP) e a doença de Huntington (DH), provocam ainda disfunção e morte precoce das células neuronais, levando perturbações no sistema nervoso humano e, por conseguinte, em efeitos devastadores na vida dos pacientes e das suas famílias.

A doença de Huntington é uma condição neurodegenerativa progressiva rara, com herança autossómica dominante, caracterizada por uma tríade de sintomas progressivos que incluem perturbações motoras, cognitivas e neuropsiquiátricas. Estes sintomas culminam, em última instância, na morte prematura do paciente. A causa subjacente da DH reside na mutação do gene (*HTT*) que codifica a proteína huntingtina (Htt), localizado no cromossoma 4. A huntingtina é expressa em vários tecidos, mas encontra-se em maior abundância no cérebro, onde desempenha várias funções essenciais. As funções associadas à huntingtina permanecem pouco claras, mas os estudos apontam para que a proteína esteja envolvida em processos celulares como o transporte intracelular, a regulação da expressão génica e a sinalização celular. Notavelmente, não é a proteína completa que desempenha um papel significativo na fisiopatologia da doença. Os fragmentos N-terminais da mesma, particularmente aqueles codificados pelo Exão 1 gerados por *splicing* aberrante, são identificados como sendo as variantes mais citotóxicas. De facto, estes fragmentos N-terminais foram detetados em tecidos cerebrais *post-mortem* de indivíduos afetados pela DH. O fragmento do Exão 1 da proteína é composto por três domínios. Primeiro, a região N17 que contém os primeiros dezassete aminoácidos da proteína, que é seguida por um domínio que contém repetições consecutivas do aminoácido glutamina (CAG), compondo uma região de poliglutaminas (poliQ) central no fragmento. Em indivíduos saudáveis, a sequência de poliQ tem um número normal de repetições de glutaminas (até 38 repetições). Na mutação que conduz ao fenótipo da doença, ocorre uma expansão anormal da sequência de poliQ da huntingtina (≥39 repetições de glutaminas), passando a existir na sua forma mutada (mHtt). Seguido da poliQ, encontra-se uma região rica em prolinas, que se pensa ser responsável por manter a proteína no seu estado solúvel. Quanto maior for a expansão da sequência de poliQ, maior o risco de desenvolver a DH e mais grave tende a ser a manifestação dos sintomas. A expansão anormal de glutaminas na huntingtina mutante confere propriedades tóxicas à proteína, pelo que esta tende a formar agregados insolúveis dentro das células cerebrais, especialmente nos neurónios, como inclusões intracelulares ou corpos de inclusão. A formação desses agregados interfere com múltiplos processos celulares, resultando em disfunção neuronal, disrupção das vias de sinalização e eventual morte das células nervosas.

A neuroinflamação é frequentemente observada em pacientes com HD, tendo início em fases precoces da patologia, antes do aparecimento dos primeiros sintomas. Os processos de neuroinflamação são desencadeados em resposta a vários estímulos, incluindo infeções, lesões traumáticas cerebrais, metabolitos tóxicos ou respostas autoimunes. No sistema nervoso central, que inclui o cérebro e a espinal medula, as células da microglia são as células imunitárias inatas residentes que são ativadas em resposta a estes estímulos. Os astrócitos são um tipo de célula glial no sistema nervoso central que desempenham um papel crucial na regulação da neuroinflamação, secretando mediadores pró-inflamatórios, nos quais

se inclui a proteína S100B. A proteína S100B, membro da família das proteínas S100 é uma pequena proteína de ligação a Ca^{2+}, com alta abundância no cérebro, representando cerca de 0.5% das proteínas do mesmo. Esta proteína é produzida principalmente por astrócitos, e desempenha um papel crítico na regulação de diversos processos celulares, tais como a proliferação e diferenciação celular. Caracteriza-se ainda pela função de regulação e manutenção da homeostase de cálcio no espaço intracelular. Níveis elevados de S100B no soro têm sido associados a várias condições neurológicas, incluindo lesões traumáticas e doenças neurodegenerativas, desencadeando um interesse crescente na investigação. Estudos recentes associaram uma dualidade na função da proteína S100B. Por um lado, em concentrações elevadas, atua como uma citocina pró-inflamatória, desempenhando um papel potencialmente prejudicial na neuroinflamação. Por outro lado, quando em concentrações fisiológicas e na presença de cálcio, a S100B apresenta funções de chaperão molecular, apresentando a capacidade de inibir a agregação de proteínas amiloidogénicas, como o péptido Aβ e a proteína tau, ambas associadas à doença de Alzheimer. Além disso, os níveis de expressão desta proteína encontram-se aumentados em astrócitos de pacientes afetados por doenças como a Doença de Parkinson e a Esclerose Lateral Amiotrófica (ELA). Foi também reportado um incremento significativo nos níveis de S100B em fluídos biológicos, como no líquido cefalorraquidiano, em indivíduos diagnosticados com Esclerose Múltipla (EM). Devido ao seu envolvimento documentado numa ampla gama de doenças neurodegenerativas, a proteína S100B tem despertado um interesse significativo na área da neurologia. A possibilidade de estender os seus efeitos como chaperão molecular a outras doenças neurodegenerativas, como a Doença de Huntington, abre novas perspetivas de investigação das interações entre a S100B e a huntingtina.

Assim, o presente estudo teve como principal objetivo investigar o potencial envolvimento da proteína S100B na agregação da huntingtina e, consequentemente, na patologia da DH. Para abordar esta questão, foram realizados ensaios celulares e ensaios *in vitro*. Numa primeira fase, purificou-se uma forma de S100B com uma Myc tag (S100B-Myc), de modo a ser utilizada em estudos celulares posteriores. Os ensaios celulares pressupuseram, numa primeira abordagem, investigar as interações entre a S100B e a huntingtina utilizando um ensaio de bio-luminescência e co-imunoprecipitação, para a deteção de interações proteína-proteína em células de mamíferos. Numa segunda abordagem, estudaram-se os efeitos exercidos pela presença de S100B endógena (utilizando o plasmídeo de S100B-Cerulean e sua transfeção) e exógena (através da adição de S100B-Myc purificada ao meio) em células HeLa vivas. Estas células expressavam duas formas distintas do Exão 1 da huntingtina: uma versão wild-type, que possui uma sequência de poliglutaminas de tamanho considerado normal, e uma versão mutante, na qual o número de repetições de glutaminas na poliQ ultrapassa o limiar patofisiológico. Esta análise foi realizada recorrendo a ensaios de fluorescência biomolecular (BiFC), com o intuito de avaliar potenciais alterações na quantidade e dinâmica das inclusões de Htt, sob a forma de grandes agressomas ou de pequenos agregados (*punctae*) na região citoplasmática, na presença da proteína S100B. Os resultados preliminares indicaram que, qualquer forma de S100B, seja endógena ou exógena, não teve um impacto estatisticamente significativo na formação de agressomas dentro das células. Por outro lado, observa-se que no que diz respeito à percentagem de células com agregados tipo punctae, em que, comparativamente ao grupo de controlo, tanto as condições com a transfeção de S100B-Cerulean e com a adição de S100B-Myc demonstram uma redução na população de células que exibem este tipo particular de estrutura, com percentagens de 77,2% e 85,9%, respetivamente, cujas diferenças não foram igualmente estatísticamente significativas. Estes resultados sugerem que existe uma tendência para uma redução no aparecimento de agregados tipo punctae na presença de S100B.

Os ensaios *in vitro* foram realizados de modo a aprofundar a compreensão de como a proteína S100B impacta a agregação da huntingtina. A abordagem utilizada para esta questão foi uma técnica de fluorescência baseada em FRET, conhecida como FRASE. Os resultados desses ensaios demonstraram de forma conclusiva que, na presença de cálcio, a S100B suprime significativamente a agregação da Htt

na sua forma mutada, operando a concentrações suprastoiquioimétricas. Com este trabalho, pretendeu-se investigar a relação entre a agregação de huntingtina e a proteína S100B, podendo este constituir uma base para o desenvolvimento de novas abordagens terapêuticas no contexto da DH.

Palavras-Chave: Doença de Huntington, huntingtina, agregação proteica, S100B.

List of figures

List of tables

Abbreviations

Aβ – Amyloid-beta peptide

AD – Alzheimer's Disease

ALS – Amyotrophic Lateral Sclerosis

ANS – 8-Anilinonaphthalene-1-sulfonic Acid

BDNF – Brain-derived neurotrophic factor

BiFC – Bimolecular Fluorescence Complementation

CD – Circular dichroism

CAG – Trinucleotide sequence in DNA for glutamine

CNS – Central Nervous System

COX-2 – Cyclooxygenase-2

CSF – Cerebrospinal Fluid

DTT – Dithiothreitol

DAMP – Damage-associated molecular pattern

ECL – Enhanced chemiluminescence

EDTA – Ethylenediaminetetraacetic Acid

FGFR1 – Basic fibroblast growth factor receptor 1

FRASE – FRET-based mHtt aggregate seeding

FRET – Förster Resonance Energy Transfer

GFP – Green Fluorescent Protein

GO – Gene ontology

GST – Glutathione S-Transferase

HAP1 – Huntingtin-associated protein 1

HAP40 – Huntingtin-associated protein 40

HD – Huntington's Disease

HEAT – Huntingtin, Elongation factor 3, PR65/A, TOR1

HEPES – 4-(2-Hydroxyethyl)piperazine-1-ethanesulfonic Acid

Htt – Huntingtin protein

HttExon1 – Exon1 of huntingtin

HTT – Huntingtin gene

iNOS – Inducible nitric oxide synthase

IPTG – Isopropyl β-D-1-thiogalactopyranoside

IT15 – Interesting transcript 15

LLPS – Liquid-liquid phase separation

LuTHy – Luminescence-based two-hybrid

mRNA - Messenger RNA

MS – Multiple Sclerosis

MPTP – 1-Methyl-4-phenyl-1,2,3,6-tetrahydropyridine

NES – Nuclear export signal

N17 – N-terminal region of huntingtin, that comprises the first 17 amino acids

ON – Overnight

OD_{600nm} – Optical density at 600 nm

PD – Parkinson's Disease

pI – Isoelectric point

PMSF – Phenylmethylsulfonyl Fluoride

polyQ – Polyglutamine tract

PP – PreScission Protease

PRD – Proline-rich domain

Put – Putamen

PTMs – Post-translational modifications

RAGE – Receptor for advanced glycation end products

ROS – Reactive oxygen species

τ – Tau protein

TFC – Total Functional Capacity scale

TH+ – Tyrosine Hydroxylase-positive

Tm – Melting temperature

Tris-HCl – Tris(hydroxymethyl)aminomethane-hydrochloride

UHDRS – Unified Huntington's Disease rating

UV – Ultraviolet

ε_{280nm} – Molar extinction coefficient at 280 nm

λ – Wavelength

Index

I. Introduction

The human nervous system is an extraordinary and distinctive attribute of our existence, highlighting the uniqueness of our species in the natural world [1]. It is a highly complex network composed of neuronal and support cells, comprising approximately 86 billion neurons that intercommunicate through a vast number of electrochemical signals in response to stimuli or sensory inputs since the first day of our birth [2] [3].

Neurons become terminally differentiated during development, meaning that they lose their ability to divide and proliferate [4], therefore they must endure for an individuals' lifetime. Due to cellular death, the age-associated decline of the number of neurons is inevitable. Neurodegenerative diseases such as Alzheimer's Disease, Parkinson's Disease and Huntington's Disease induce dysfunction and precocious neuronal cell death, leading to impairments in the human nervous system and consequently to devastating effects on patients and their families' lives.

The introduction to this thesis is divided into three sections. The first section provides an overview of Huntington's disease, the neurodegenerative condition that is the central focus of this work. It delves into the discovery, clinical manifestations, etiology, and epidemiology of Huntington's disease. In the second section, we introduce two key proteins that play roles in neuronal dysfunction associated with Huntington's disease. One of these proteins is the well-known Huntington protein, which is the primary culprit implicated in the disease's pathology, forming deposits that are closely linked to neuropathological events. The other protein is S100B, a member of the S100 protein family, known for its multifunctional properties. Within this family, S100B stands out as a protein with a novel function – it acts as an aggregation chaperone. This introduction sets the stage for the objectives of this thesis, where we will investigate the relationship between the aggregation of Huntington's protein and the chaperone function of S100B.

1.1. Huntington's Disease (HD)

1.1.1. Discovery

The first medical descriptive account of the disorder was dated in 1872, when the young doctor George Huntington wrote an article "On chorea", where he presented the clinical features of chorea and also an elucidation of a particular disease that seemed "to obey certain fixed laws" as a "medical curiosity", that was after named "Huntington's chorea" and that is known since the 1960s until today as Huntington's Disease. He described its "hereditary nature", the onset occurring at "adult or middle life", the muscular atrophy with "irregular and spasmodic action", a predisposition for the development of psychiatric symptoms and a steady progression of the disorder without any signal of remission, stating that "Once it begins, it clings to the bitter end" [5] [6]. One hundred and fifty years after these findings, the improved understanding of the pathophysiology and the clinical manifestations of the disease have led to deepen the studies to identify the genetic mechanisms underlying the disorder and exploring potential therapeutic approaches that can directly inhibit or slow its progression.

1.1.2. Clinical Manifestation of HD

Huntington's disease (HD) is a rare progressive neurodegenerative illness with autosomal dominant inheritance that manifests with a triad of symptoms, leading to motor, cognitive and neuropsychiatric disturbances, that ultimately lead to the patient's premature death [7].

The gradual progression of the disease makes it difficult to diagnose before severe and irreversible symptoms appear. The course of HD can be divided into two periods, "premanifest" and "manifest". The premanifest stage is clinically silent, therefore the patient shows no signs of clinical symptoms. This stage can be further subdivided into two periods: presymptomatic and prodromal.

In the presymptomatic period, individuals are not clinically distinguishable from controls, and this goes from 10-15 years before onset [8]. The prodromal period is characterized by early and subtle changes in motor function, cognition, and behaviour. Once motor and cognitive signs begin, they progress steadily over the course of the disease, until the manifest stage is reached and becomes possible to establish a clinical diagnosis of Huntington's disease [9] **(Figure 1.1)**.

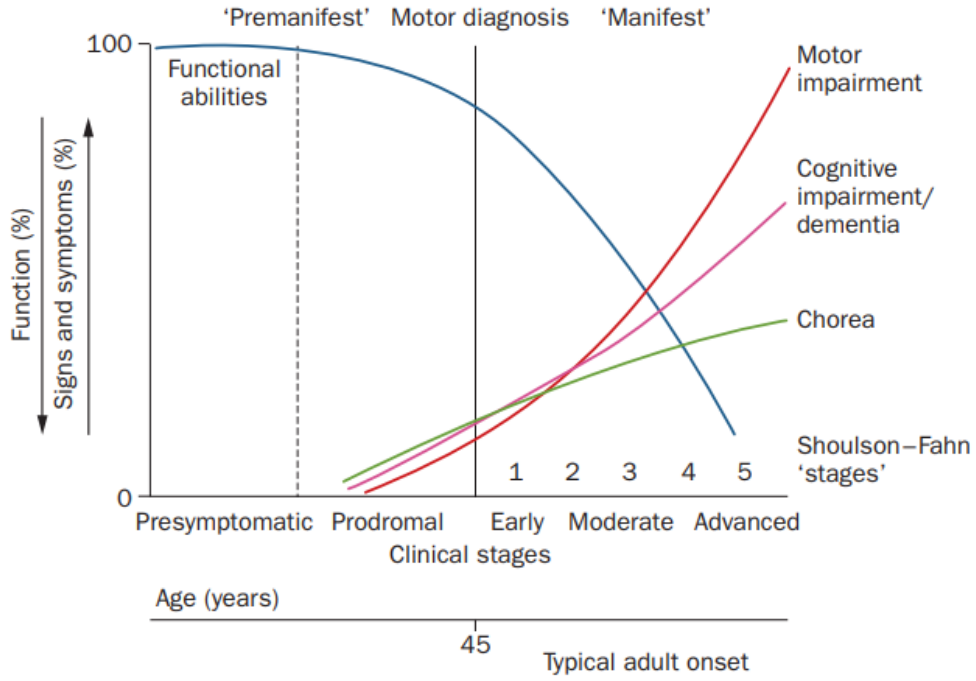

Figure 1.1. – Natural history of clinical HD. The term "premanifest" is the period prior to the manifestation of diagnosable signs and symptoms of HD. During the "presymptomatic" period, there are no signs or symptoms present. In the "prodromal" stage, early and subtle signs and symptoms are present. Manifest HD is characterized by slow progression of motor and cognitive abnormalities, with chorea often prominent early but stabilizing or even decreasing later. Other finer motor impairments such as incoordination, bradykinesia and rigidity progress more steadily. (Adapted from [8])

As the disease progresses, symptoms worsen throughout the time, resulting in different stages of disease severity. To assess the severity of HD and evaluate functional abilities, both Shoulson-Fahn functional capacity rating scale and Unified Huntington's Disease Rating scale (UHDRS) are used. The Shoulson-Fahn scale was proposed in 1979 and assesses independence of individuals in daily tasks, using a Total Functional Capacity (TFC) scale. The scale includes 5 categories, graded from 0 to 2 or 3, with 0 representing the poorest function, assessing various aspects of daily living, such as managing personal finances, level of care, domestic chores, capacity to perform activities in daily living (ADL) and engagement in occupation. These stages appear represented in **Table 1.1**. Functional capacity is ranked

from stage 1 to stage 5, with stage 1 representing the most independent level of function. These stages are also represented in **Table 1.2**.

The UHDRS was published in 1996 by the Huntington Disease Group and is the most commonly used clinical rating scale for assessment and overall measure of disease burden. Different domains of clinical performance and capacity in HD are assessed, including motor function, cognitive function, behavioural assessment, functional capacity evaluation, independence scale and functional capacity (TFC) [11]. The clinical assessment of premanifest individuals and individuals that have familial history of HD is based on a "diagnostic confidence score", which scores motor signs according to the clinician's evaluation of how the motor abnormalities represent HD, and it is rated from 0 (normal, with no motor abnormalities suggestive of HD) to 4 (motor abnormalities that are unequivocal signs of Huntington's disease with ≥99% confidence) [12].

Table 1.1 – Total Functional Capacity rating scale for HD proposed by Shoun and Fahn and incorporated by UHDRS. The scale includes 5 categories – occupation, finances, activities of daily living (ADL), domestic chores and care level – graded from 0 to 2 or 3, with 0 representing the lowest level of function. [13] [11]

Total Functional Capacity rating scale		
Domain	**Ability**	**Score**
Occupation	Unable	0
	Marginal work only	1
	Reduced capacity for usual job	2
	Normal	3
Finances	Unable	0
	Major assistance	1
	Slight assistance	2
	Normal	3
Activities of Daily Living	Unable	0
	Minimal impairment	1
	Normal	2
Domestic Chores	Total care	0
	Gross tasks only	1
	Minimal impairment	2
	Normal	3
Care level	Full-time nursing care	0
	Home for chronic care	1
	Home	2
TOTAL	Range 0 - 13	

Table 1.2 – TFC total score and respective stages of disease progression. TFC total score and respective stages of disease progression. Functional capacity is ranked from stage 1 to stage 5, with stage 1 representing the most independent level of function. [13]

TFC Total Score	Stage
11-13	I
7-10	II
3-6	III
1-2	IV
0	V

1.1.3. Aetiology

The groundbreaking discovery of the genetic defect underlying HD was assigned in 1983, marking it as the first disease-associated gene to be molecularly mapped using only DNA markers. Researchers successfully identified it residing within the chromosome 4, paving a great insight on the genetic basis of the disease. The significance of this milestone extended far beyond HD, as it inspired similar studies in numerous other disorders and played a pivotal role in shaping the concept of the Human Genome Project [14]. This pioneering effort ultimately led to the remarkable discovery of the HD-associated *HTT* gene, its specific mutation, and the identification of its unique protein product, ten years later. In 1993, a group of researchers of the Huntington's Disease Collaborative Research Group identified the DNA sequence and the precise nature of the HD-associated mutation in the *IT15* (interesting transcript 15) gene, located on the short arm of chromosome 4 (4.p16.3). The *IT15* gene encoded a previously undescribed protein, with a molecular weight of 348 kDa, and the most outstanding observation reported was that the reading frame of this gene contained a polymorphic trinucleotide repeat region consisting of C (cytosine), A (adenine) and G (guanine) nucleotides – encoding for a polyglutamine tract (polyQ). Analysis of the HD-associated genetic region demonstrated that non-HD control individuals typically had a range of 6 to 35 repeats in the CAG polyglutamine stretch, while individuals affected by HD exhibited 40 or more CAG repeats in the same genetic locus [15]. Therefore, it was inferred that HD is caused by the expansion of the unstable CAG repeat in the *IT15* gene, that was later designated as huntingtin gene (*HTT*).

Subsequent studies further elucidated the role of the CAG repeat length in the onset and progression of the disease. It was verified that individuals with below than 27 CAG repeats belong to the normal range phenotype, which consequently, are not expected to have any symptoms or are at risk of developing the disease. The intermediate range, that goes from 28 to 35 repeats, is considered nonpathological, also having no direct pathophysiological phenotypic consequences on the individual. On the other hand, once the CAG length surpasses a threshold of 39 repeats, a significant shift occurs, as individuals develop HD with full penetrance [10].

Although an intermediate range does not correlate with the appearance of the disease on the carrying individual, it is important to note that CAG repeats are significantly instable and may further increase its length when transmitted to subsequent generations – phenomenon known in HD and other diseases as paternal anticipation - as sperm from males shows greater repeat variability and larger repeat sizes than somatic tissues, which might happen during spermatogenesis [16]. As a result, a father with an intermediate repeat length, although not affected himself, may potentially have a child with an expanded glutamine stretch, leading to the appearance of a pathological phenotype in his offspring [10] [17] [18]. This is notable because, although most HD patients display disease onset in mid-life, the rare individuals

with juvenile-onset HD result from paternal inheritance [19]. Juvenile Huntington's disease is a less common, early-onset form of Huntington disease that begins in childhood or adolescence and accounts for 5-10% of all HD cases [20]. The explanation for an early onset of the disorder is that the age of HD diagnosis, primarily determined by motor symptoms, is largely influenced by the specific length of the expanded CAG tract, existing a notable inverse correlation between the length of CAG repeats and the age of onset in HD. This means that longer CAG repeats are associated with an earlier manifestation of clinical symptoms, as graphically depicted in **Figure 1.2**.

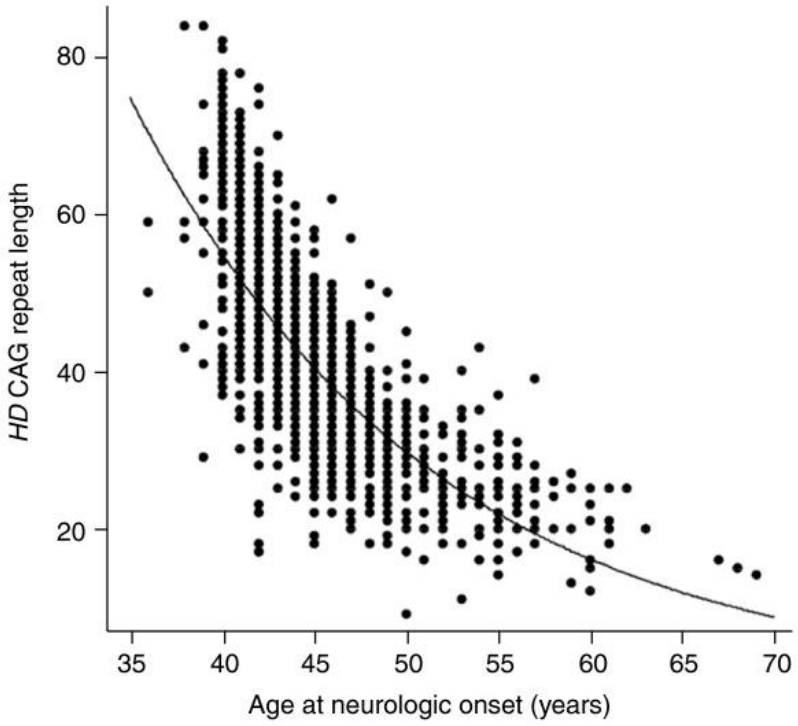

Figure 1.2 – Inverse correlation of age at neurologic onset and CAG repeat length. The plot shows the CAG repeat length of 1200 HD patients measured in blood DNA (x-axis) against age at neurologic onset (y-axis). The black line represents the best-fit simple logarithmic regression to the data points. The length of the CAG repeats accounts for approximately 67% of the overall variation in age at neurologic onset. [21]

Although the primary determinant of if and when an individual will develop HD are the presence and length, respectively, of the expanded CAG stretch, it is important to note that the CAG repeat length alone can only partially account for the variability in age of onset. Patients with the same CAG repeat length may experience different psychiatric, cognitive and peripheral phenotypes, as well as distinct progression of symptoms or suffer death at different times after disease onset. This suggests that additional factors beyond the CAG repeat length, including environmental factors, such as lifestyle, diet, and exposure to toxins, can potentially impact the age of onset and progression of HD [22]. Additionally, genetic modifiers have been identified as influencing the disease course. These genetic modifiers can influence the CAG repeat length and modify the phenotype of HD, leading to differences in symptoms and disease progression [19]. The complex interplay between CAG repeat length, environmental factors and genetic modifiers is crucial for comprehending the variability observed in HD.

1.1.4. Epidemiology

The epidemiology of HD depends on identifying individuals with the expanded polyglutamine mutation, and on thorough examination of clinical signs to assess disease onset. A recent systematic review and meta-analysis reported a global prevalence of 3,92 per 100,000, whereas the incidence of HD is estimated to be 0.47 per 100,000 yearly. The incidence is calculated by taking the ratio of the number of new HD cases diagnosed in a given year to the total number of patient-years at risk. The studies mentioned were combined with studies from a meta-analysis in previous reports, in a total time period between 1985-2022. [23] [24]

1.2. Huntingtin

Huntington's disease arises from an elongated repetition of CAG sequences within the *HTT* gene, which is responsible for encoding the huntingtin protein (Htt). This expanded CAG repeat gives rise to a mutated form of the Htt protein, triggering a cascade of events that lead to neurodegeneration and the development of the disease's associated distressing symptoms.

1.2.1. Structure and function

Huntingtin is a large protein encoded by the *HTT* gene, composed of 3144 amino acids, and has a molecular weight of approximately 348 kDa. It is ubiquitously expressed throughout the body though observed at highest levels in the brain, particularly in neurons within the CNS, and on the male gonads [25]. Despite its size, wild-type huntingtin is soluble and in cells it is primarily present in the cytoplasm, although it can also appear in the nucleus [26] [27] . Intracellularly, the protein is present in association with a multitude of organelles, including the nucleus, endoplasmic reticulum, and Golgi complex. It is also found in neurites and at synapses, where it forms associations with vesicular structures such as endosomal compartments or caveolae, clathrin-coated vesicles and microtubules. Due to its widespread subcellular localization, the precise function and role of huntingtin is still not clear. [25]

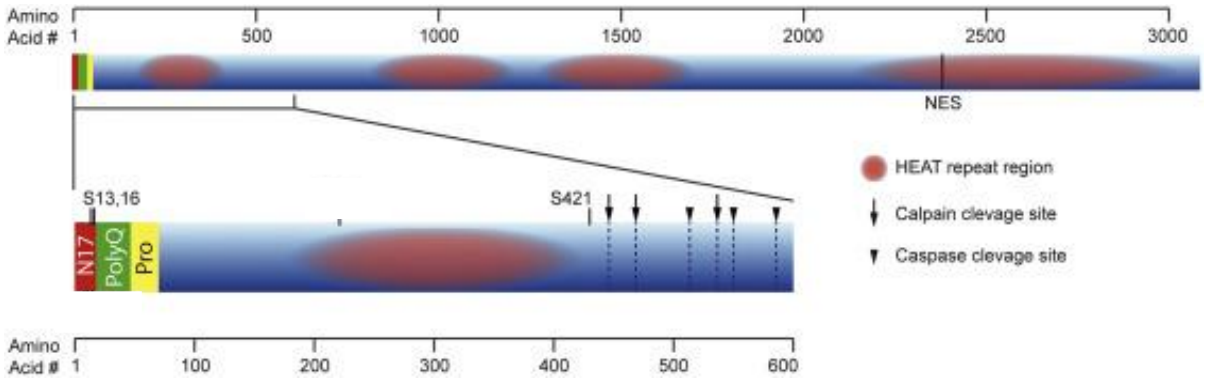

Figure 1.3 – Structural overview of the huntingtin protein. Sites of calpain and caspase cleavage are shown with a longer and shorter arrow, respectively. Clusters of HEAT repeats are denoted in red. N17, first seventeen aminoacids in the N-terminus; Pro, proline-rich region; NES, nuclear export signal. (Adapted from [26])

Structurally, Htt is an α-helical protein that has several distinct domains, which boundaries and activities are still not fully understood. The N-terminal domain is sensitive to proteases, such as caspases and calpains in specific positions. Posttranslational modifications can have profound effects on the biological and biophysical properties of proteins associated with misfolding and aggregation. Htt may undergo different PTMs, including phosphorylation, SUMOylation, lipidation, ubiquitination, acetylation, proteolytic cleavage, and palmitoylation. In the presence of the HD mutation, some PTMs are significantly altered and can result in changes in the clinical phenotype [28]. For example, recent studies showed that phosphomimetic mutations at specific residues (S13/S16) within the N17 region (first 17 aminoacids of the protein in the N-terminal) can modulate the subcellular localization, stability, aggregation, and toxicity properties of mutant Htt in different preclinical models [29]. In addition, phosphorylation of S421, has been shown to have a protective role against neuronal toxicity [30] **(Figure 1.3)**. These, among several other studies, are raising interest in these PTMs having a crucial role as possible modifiers of the biological, biophysical, and biochemical properties of mHtt.

Within the N-terminus, there is a polyglutamine domain, which begins at the eighteenth amino acid of the protein, and as mentioned before, in unaffected individuals, contains up to 34 CAG repeats. The polyglutamine tract had been object of intense study to understand how the folding of this structure was. In 1994, Perutz *et al.* observed that the polyQ tract had a polar zipper like structure, and further studies suggested that this portion of the protein is a key regulator of protein-protein interactions between huntingtin and several partners. Downstream of the polyQ there is the proline-rich-region that is responsible for maintaining the protein in a soluble state, and this domain is followed by the presence multiple 50 amino acid-long HEAT (Huntingtin, Elongation factor 3, protein phosphatase 2A, the yeast kinase TOR1) domains, that work as scaffolds for protein-protein interaction [31]. Huntingtin also contains a functionally active carboxy (C)-terminal nuclear export signal (NES) sequence that allows the protein to exit the nucleus and relocate to the cytoplasm [32], which suggests that huntingtin, or a specific part of the protein, may play a role in the transportation of molecules from the nucleus to the cytoplasm. Given its presence in both the nucleus and cytoplasm, it is implied that the protein is dynamic between these two compartments. [33]

1.2.1.1. 3D structure

Extensive research has been conducted throughout the years to unravel the three-dimensional structure of Htt at high resolution, which have been hindered by its high flexibility [34]. To overcome this, a recent cryo-electron microscopy (cryo-EM) uncovered the 3D structural arrangement of the protein in a complex with HAP40 (Huntingtin Associated Protein 40) that has been found to be an interactor with huntingtin in cells. Subsequent investigations further improved the obtained resolution from 4 Å to 2.6 Å [34] [35]. Huntingtin in connection with HAP40 is stabilized in a particular conformation, enabling to derive its 3D structure. [35]

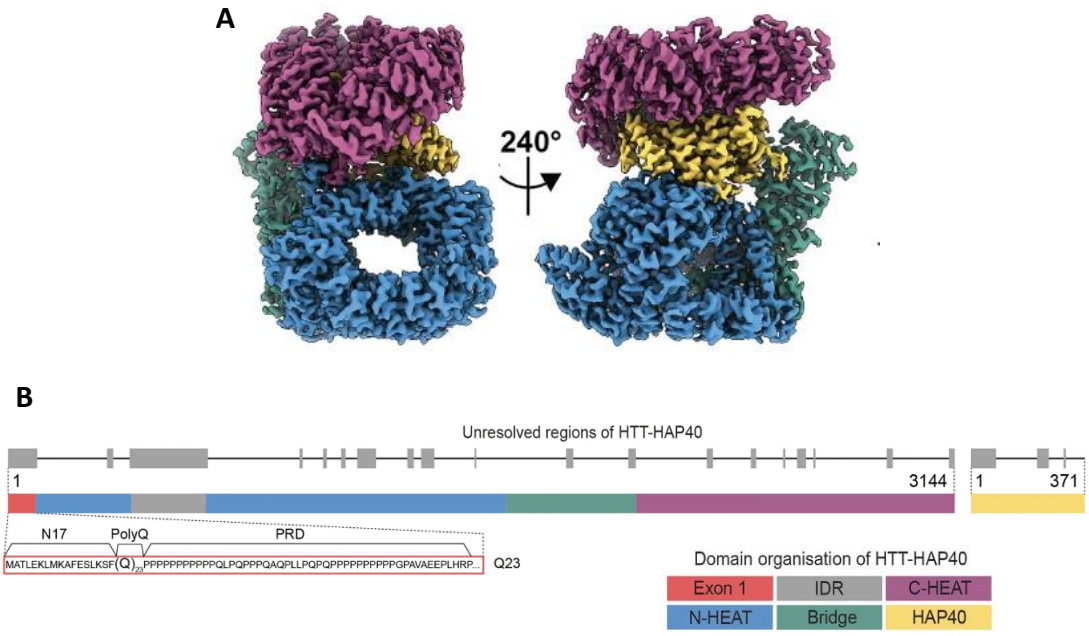

Figure 1.4 – 3D structure of Huntingtin in complex with the interactor HAP40, that stabilises Htt structure. (A) Cryo-EM volume of Htt-HAP40 resolved to 2.6 Å with Htt N-HEAT in blue, bridge domain in green, C-HEAT in purple and HAP40 in yellow. **(B)** Domain organisation of Htt shown mapped to linear sequence. Unresolved regions of the structure are in grey, and the construct used comprising wild-type Htt with 23 repeats of glutamines (Q23) is detailed, as well as the colour code for all regions represented. (Adapted from [35])

1.2.1.2. Function

Although the precise function of huntingtin is not fully uncovered or understood, investigations have provided insights into the roles associated with the protein in distinct cellular processes.

Huntingtin is essential not only for early development, as its depletion is lethal in mice embryos [36], but also in young and adult mice, where its ablation causes the development of acute pancreatitis and alterations in brain homeostasis, respectively [37].

In addition, huntingtin is involved in the control of axonal and vesicular transport [25]. A large number of Htt interactors function in microtubule-based axon trafficking. Huntingtin- associated protein (HAP1) has a role on mediating the interaction between Htt and microtubule motor proteins, kinesin, dynactin and dynein, and also their associated co-factors. Kinesin and dynein are related to antero- and retrograde axon transport, suggesting that huntingtin is involved in both processes. In Htt knock-out mice, vesicular and mitochondrial trafficking also appear impaired, and progressive brain degeneration is also observed [38]. As mentioned before, Htt has been found to be associated with vesicles as it interacts with HIP1,

a clathrin-binding protein, to mediate endocytosis, also being involved in the trafficking of materials into the cell [39] [40]. It also interacts with β-tubulin in the microtubules, indicating a potential role in cytoskeletal anchoring. [41]

There is also evidence that huntingtin upregulates the expression of brain-derived neurotrophic factor (BDNF) at the transcription level, a pro-survival factor produced by cortical neurons that is necessary for survival of striatal neurons in the brain [42].

It has also been observed that huntingtin has neuroprotective properties. In conditional Htt knock-out mice, the reduction of huntingtin was shown to lead to apoptotic cell death in the striatum, cortex, and hippocampus, evidencing its anti-apoptotic activity. Additionally, it has also been reported that Htt protects neurons from excitotoxicity [38].

While the exact mechanisms and functions of huntingtin are still being investigated, these findings suggest that huntingtin is involved in various cellular processes. Understanding the functions of huntingtin is important for unravelling its role in normal brain function and its implications in Huntington's disease. A recent study published in the Journal of Huntington's Disease [43] compiled data of Htt interactors from distinct studies and found that both mutant and wild-type Htt interact with more than 2,971 proteins, showing that both mutant and wild-type Htt protein interactors cluster into twelve gene ontology (GO) biological functions (**Figure 1.5**). The twelve clusters indicated functions of: 1) protein modification, 2) chromatin organization, 3) RNA splicing, 4) membrane trafficking, 5) signal transduction, 6) mitochondria, 7) granule membrane, 8) macroautophagy, 9) cytoplasmic vesicle, 10) clathrin-mediated endocytosis, 11) ion channel transport, and 12) translation, some of which have already been described. Although huntingtin's function remains unknown, involvement in intracellular transport, autophagy, transcription, mitochondrial function, and signal transduction have been postulated based on its protein–protein interactions. [43]

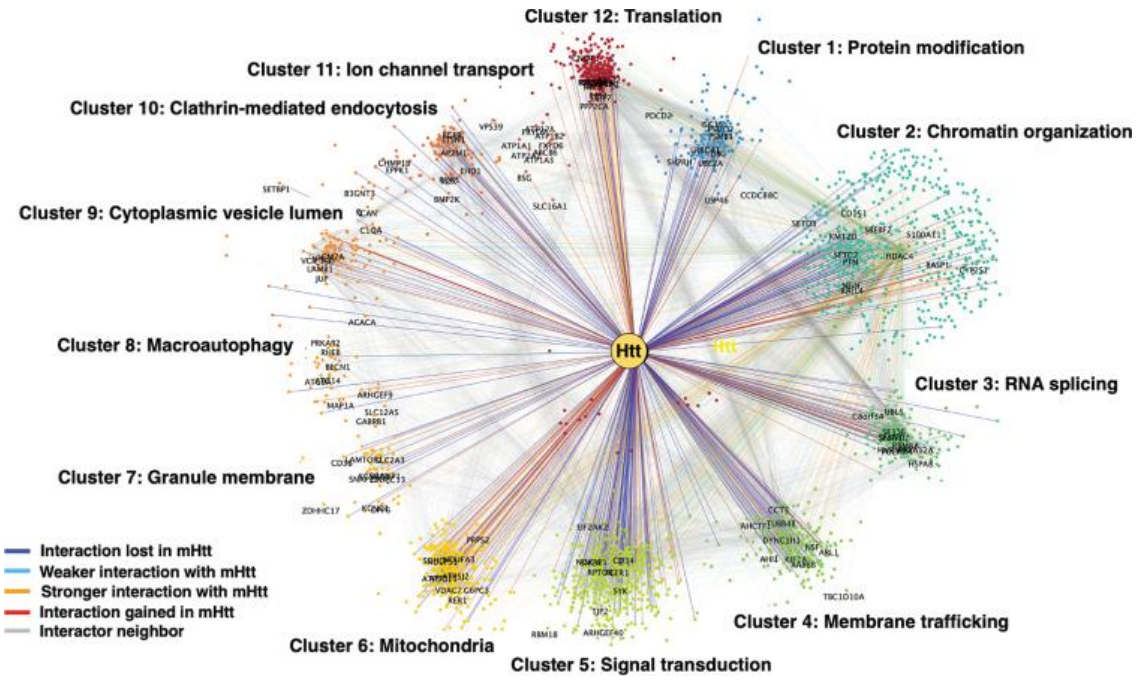

Figure 1.5 – Clustering analysis of huntingtin interactors by gene ontology (GO). Htt interactors were analysed for shared functional categories or "clusters", within different cellular molecular pathways. The resulting cluster networks are depicted in a color-coded format (see key in figure) to indicate interactions that were lost by mutant Htt, weaker with mHtt, stronger with mHtt, gained interaction with mHtt, or represented a neighbour. [43]

1.2.2. Mutant huntingtin

The development of HD is primarily driven by the mutational expansion of the polyQ tract of huntingtin, as previously mentioned. Interestingly, it is not the entire protein that plays a role in the pathophysiology, but rather the N-terminal fragments, especially those encoded by exon 1, generated by aberrant splicing, are recognized as the most cytotoxic variants. The N-terminal fragments were found in postmortem brain tissues of individuals with HD, which had led to the hypothesis that N-terminal peptides may be critical modulators of disease pathology. This discovery raised considerable interest in studying potential therapeutic strategies that involve targeting abnormal splicing or proteolytic processing of the protein [44] [45].

The accumulation of N-terminal fragments of mHtt is observed in nuclear and cytoplasmatic inclusions, which when overexpressed in mice models, leads to age-dependent accumulation and aggregation in the striatum and cortex, leading to HD-like phenotypes. [31]

The N-terminal region of exon 1 encodes the first 17 amino acids of the huntingtin protein, which are usually referred to as the N17 domain. When the mutation is present and the polyQ tract is expanded, it can lead to an aberrant splicing of the mRNA or to protease-mediated cleavage of the protein causing the formation of a N-terminal fragment of the protein. This leads to the incomplete pre-mRNA splicing of exon 1 and exon 2, and because the resulting mRNA has a stop codon after the exon 1, only this region is translated and consequently produced. [31]

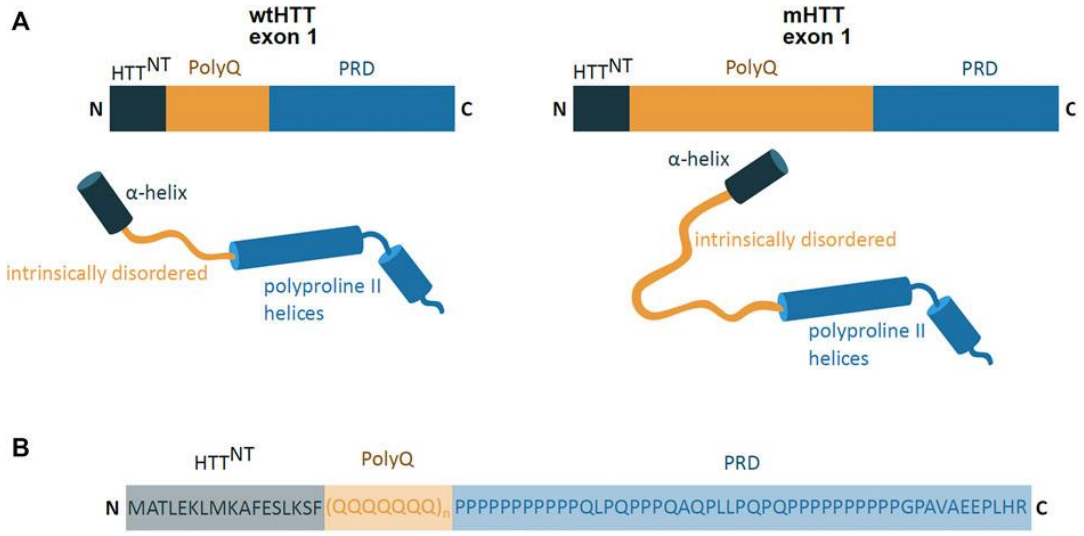

Figure 1.6 – Mutant Htt Exon 1 structure and sequence. (A) Schematic presentation of primary and secondary structure of wild-type Htt (wtHtt) and mutant Htt (mHtt) exon 1, showing the N17 (dark blue; α-helix), polyQ domain (yellow; intrinsically disordered) and proline-rich domain (PRD) (lighter blue; polyproline II helices). Mutant huntingtin contains an expanded polyQ domain compared to wild-type Htt. N and C indicate the protein N- and C-terminus. **(B)** The amino acid sequence of the exon 1 mHtt protein. Highlighted are the identified regions of N17, PolyQ and PRD (proline rich domain). The protein sequence is represented by one letter amino acid code. (Adapted from [31])

The exon 1 of the mHtt protein is composed of three domains. First, the N17 region, that presents a compact α-helical coil structure, which is thought to act as an anchor tethering the protein to intracellular membranes. The N17 region is followed by a central polyQ domain, that is a conformationally versatile peptide which can take on different structures, such as intrinsically disordered, an α-helical conformation near the N-terminal domain, or a β-sheet conformation found in mHtt aggregates. On the C-terminal of exon 1 is a 50-residues proline-rich domain, that adopts an intrinsically disordered conformation with polyproline II helices (**Figure 1.6**). The different conformations that the polyQ can adopt are influenced

by the conformational features of the flanking N- and C-terminals, as the N17 and the proline-rich region/domain (PRD), promote and inhibit the aggregation propensity of the protein, respectively, by poorly understood mechanisms. [31]

1.2.3. Protein aggregation

Protein aggregation refers to the process by which intrinsically disordered or misfolded proteins self-assemble, leading to the formation of insoluble aggregates either intra or extracellularly [46]. Protein aggregation can be classified into two general categories: amyloid or amorphous. The amyloid state involves the formation of highly structured, insoluble and fibrillar deposits, usually comprising numerous repetitions of the same protein. These fibrils are insoluble fibrillar structures formed by the assembly of proteins and peptides into a cross-β structure [47] . This particular type of structures are implicated in a wide variety of neurodegenerative diseases, such as Alzheimer's Disease, Parkinson's Disease and Huntington's Disease. On the other hand, amorphous protein aggregation is characterized by the disordered aggregation of proteins, in which individual proteins are not generally associated to any diseases when aggregated. Protein misfolding that leads to aggregation can arise as a consequence of aging, environmental stress, chemical changes, destabilizing mutations, or the absence of necessary assembly partners. Newly synthetized proteins have higher susceptibility to misfolding events and protein aggregation is thought to be toxic, particularly when the proteostasis network is impaired. [48]

1.2.3.1. Aggregation of mHtt Exon1

The pathophysiological hallmark of HD is the formation of mutant huntingtin aggregates and inclusion bodies inside cells. The formation of mHtt aggregates is self-initiated and is dependent on the polyQ region length, protein concentration and time, meaning that the mHtt is capable of aggregate spontaneously and form distinct structures such as monomers, oligomers, and fibril-rich inclusion bodies [49]. The aggregates with an oligomeric nature, observed both in the cytoplasm and on the nucleus, are small and soluble. In contrast, inclusion bodies, that can also be found in the cytoplasm and/or in the nucleus, consist of highly insoluble aggregates that are predominantly spherical and composed of fibrils, which are elongated rod-like structures formed by tightly packed mHtt proteins. The mHtt fibrils are amyloid-like structures where the mHtt polyQ region forms a tightly packed, highly rigid and dehydrated anti-parallel β-sheet core flanked by mHtt N- and C-terminal domains (**Figure 1.7**) [50]. The molecular architecture of the deposits shares several features with aggregates found in Alzheimer's and Parkinson's disease [51]. It is important to highlight that the structural information mentioned is obtained from artificial systems, such as recombinantly expressed proteins. Indeed, when examining fibrils derived from primary materials in neurodegenerative diseases like Alzheimer's and Parkinson's disease, their structures can differ significantly from those produced *in vitro*. Therefore, the extent to which artificial Htt fibrils produced from recombinantly produced proteins adequately represent fibrils generated *in vivo* is still a matter of debate [31].

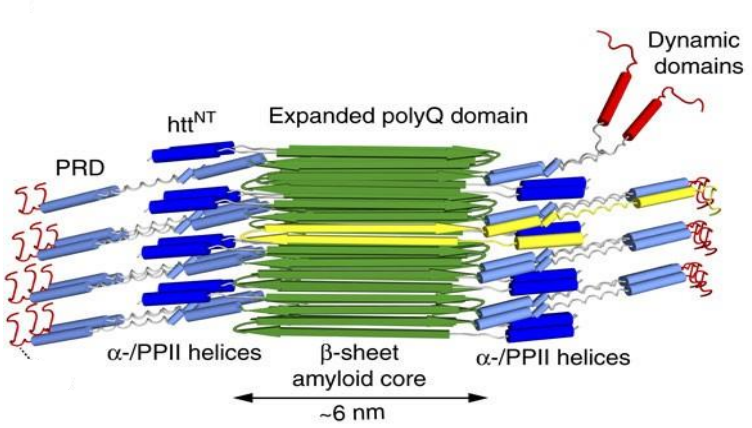

Figure 1.7 – Proposed model of HttExon1 fibrils architecture. The N17 α -helices (dark blue) and PRD poly-L-proline type II helices (light blue) are immobilized and tightly clustered around the rigid amyloid core (green β-strands). C-terminal domains show increased dynamics and flexibility, either in the form of the unstructured C-terminal tail or a subpopulation of more exposed PRDs (top right; red). An individual protein monomer with its β-hairpin-based polyQ core is shown in yellow. (Adapted from [50]).

Studies have shown that HttExon1 fragments with pathogenic polyglutamine tracts spontaneously self-assemble into insoluble aggregates with fibrillar morphology [52]. One of the models for the process of mHtt aggregation can be divided in three main stages: an initial lag phase, followed by an exponential growth phase and then a plateau. The aggregation of mHtt starts with a spontaneous nucleus inducing conformational changes of other proteins, leading to their addition to the aggregate through templated aggregation – phenomenon known as primary nucleation. [52]. This nucleus, which can be as small as a single molecule (monomer) or a few molecules grouped together (oligomer), serves as a seed for further aggregation. The expansion of the polyQ region induces conformational changes on the N-terminal domain of mHtt, which accelerates nucleation and aggregate formation [53] [54] [51]. There is also evidence of inter-neuronal transmission of HD-associated protein deposits in HD patients as the disease progresses [51]. Through the interactions established between the polyQ repeat regions, which adopt a β-sheet conformation, aggregates expand and elongate into fibrils. The growth process speeds up with the recruitment of new monomers, and these newly formed elongating fibrils are capable of inducing fibril-dependent nucleation, leading to the formation of fibril branches. Fibrillary branches are seen in Htt aggregates, but not on other amyloid fibrils [51]. Notably, wild-type huntingtin, which does not adopt fibrillary structures, can also be recruited by the elongating fibrils [53] [55].

Additionally, N-terminal HttExon1 protein can form reversible liquid-like assemblies, a process driven by the polyQ tract and the proline-rich region. A recent study showed that the protein exists in distinct liquid-like and solid-like forms, which display distinct structures. The liquid-like assemblies convert into solid-like assemblies *in vitro* and in cells. In cells, HttExon1 with sub-toxic polyQ lengths can form liquid-like assemblies, uncapable to convert to irreversible solid-like assemblies. The conversion occurs only when polyQ length extends beyond the threshold for HD. In both yeast and mammalian cells models tested, HttExon1 with different lengths of polyQ tract can all form liquid-like assemblies, but a larger polyQ length leads to the formation of more solid-like structures. [56]

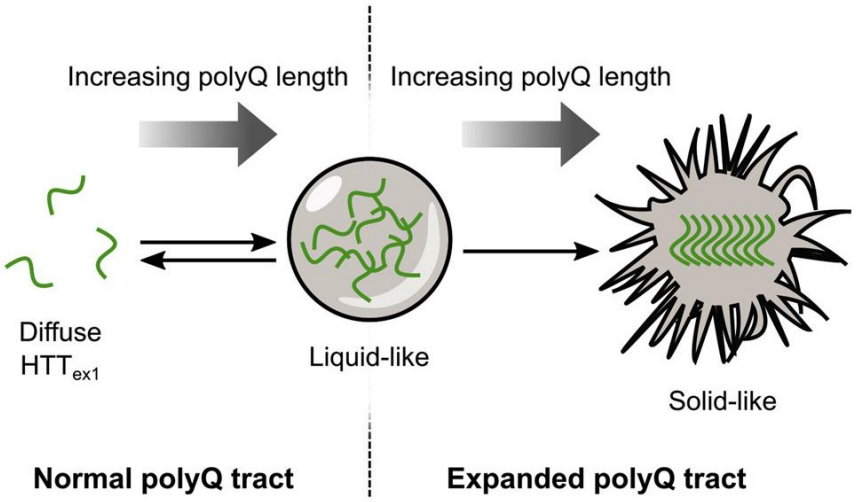

Figure 1.8 – Representation of liquid to solid phase transition underlying pathological Huntingtin Exon1 aggregation. HttExon1 with sub-toxic polyQ lengths can form liquid-like assemblies, but these do not convert to irreversible solid-like assemblies. The conversion occurs only when polyQ length extends beyond the threshold for HD. **[56]**

The research suggested that dim assemblies (liquid-like assemblies) are highly reversible structures, maintained by weak hydrophobic interactions, whereas "bright" assemblies (solid-like assemblies) are maintained by stronger, possibly amyloid-like, interactions. Authors also noted that, the rapid reversibility, circularity, coalescence, and internal mobility of dim assemblies are consistent with liquid-like properties. [56]

1.2.4. Neuropathology

Despite the ubiquitous expression of huntingtin and its mutated form, the striatum (caudate nucleus and putamen) suffers for the most substantial cell loss and decreased volume throughout the course of the disease. In **Figure 1.9**, the difference between a post-mortem brain slice at the level of the striatum of a control individual and a HD individual brain is clear. There is significant loss of brain volume, especially in the caudate nucleus and the putamen.

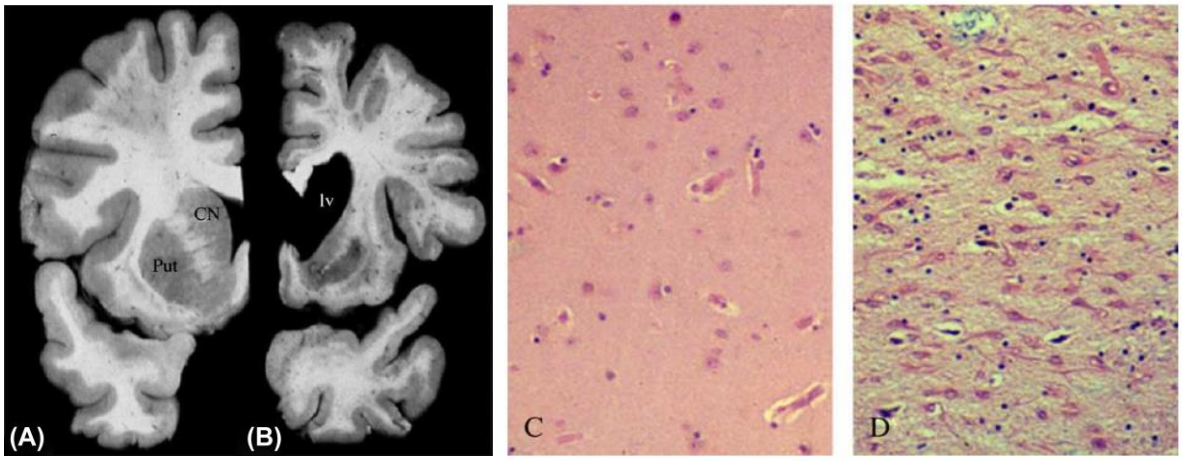

Figure 1.9 – Macroscopic and microscopic neuropathology in Huntington's Disease. (A) Coronal section through the brain of a healthy control subject and (B) a 62-year-old female with Huntington disease (HD) **(B)** Stark atrophy of the caudate nucleus (CNs) and putamen (Put) as well as the enlarged lateral ventricle (lv) in the HD brain. **(C)** The corresponding hematoxylin and eosin staining of the caudate nucleus shows significant neuronal loss and astrogliosis in Huntington's disease **(D)**, in comparison with the normal control (C). (Adapted from [57])

While other brain regions are involved, the most striking neuropathological changes are gross atrophy of the neostriatal nuclei, the caudate nucleus, and putamen, with concomitant marked neuronal loss, astrogliosis and brain shrinkage [57]. Not all neurons are affected equally, there is specific loss of efferent medium spiny neurons [58], showing a selective pattern of neuronal vulnerability and topographic susceptibility to the mutant protein.

1.3. S100 proteins

The members of the S100 protein family are multifunctional proteins with a regulatory role in a variety of cellular processes. They are calcium-binding proteins, some of which are classified as Ca^{2+} sensors, consisting in the largest subgroup of the Ca^{2+}-binding EF-hand protein group. They were first identified by B.W. Moore in 1965, and the name "S100" is attributed to the fact that these proteins are soluble in a 100%-saturated ammonium sulphate solution at neutral pH [59]. Although S100 proteins actions are exerted usually through calcium-binding, Zn^{2+} and Cu^{2+} ions have also been shown to regulate the biological activity of some members of this family [60]. S100 proteins are found exclusively in vertebrates, with expression profiles indicating that the more complex the vertebrate organism, such as in mammals, the more diverse the expression of S100 proteins, pointing toward specified functions of S100 proteins in the development of the organism. The individual members of the S100 family show a tissue- and cell type -specific expression pattern, with each protein having a unique expression and distribution profile amongst different tissues and cell types. S100 proteins have high expression levels in the brain and heart across all organisms [61]. The diverse expression patterns within the S100 protein family parallel the broader significance of calcium ions (Ca^{2+}) in cellular processes. Ca^{2+} can function as a second messenger and is involved in the control of a variety of cellular processes, ranging from muscle contraction to cell differentiation or even cell death. The regulation of intracellular Ca^{2+} levels and Ca^{2+} signalling is assured by several Ca^{2+} transporters and membrane channels. Notably, Ca^{2+}-binding proteins share a common ancestor, and therefore also share the ability to regulate the intracellular Ca^{2+} levels and many Ca^{2+}-signalling pathways.

While some EF-hand proteins have been implicated in maintaining intracellular Ca^{2+} homeostasis, the majority of these proteins are thought to transmit Ca^{2+} signals by binding to, and thereby regulating, specific target proteins when activated in their Ca^{2+} bound conformation [62] [63]. Within the EF-hand superfamily, the S100 proteins form the largest subfamily, with 25 currently known members comprising a set of non-ubiquitous Ca^{2+}-modulated proteins implicated in multiple intracellular and/or extracellular regulatory activities [62]. The family of EF-hand proteins can be divided into two primary classes: Ca^{2+} sensors, which transduce the Ca^{2+} signals, and Ca^{2+} signal modulators, which are involved in uptake, transport, and buffering of the Ca^{2+} signal. The distinction between these two groups is correlated with the $Ca2^{+}$-binding-induced conformational changes. In general, Ca^{2+} sensors undergo significant conformational changes between the Ca^{2+}-free (apo state) and the Ca^{2+}-bound state, most often exposing hydrophobic patches required for the interaction with specific ligands. Most members of the S100 family belong to the class of Ca^{2+} sensors. [64]

Over the past 15 years, 3D structures of S100 proteins have been revealed in three different forms: bound to Ca^{2+}, bound to its target protein, or in its apo state. Upon Ca^{2+}-binding, the proteins undergo large conformational changes that expose a hydrophobic surface responsible for target protein binding and regulation. Additionally, each S100 protein adopts a dimeric configuration, characterized by a symmetrical shape, existing mostly as homodimers (except for calbindin). Each S100 subunit or monomer is composed of two helix-loop-helix Ca^{2+}-binding domains connected by a central hinge region. The N-terminal EF-hand comprises helix I, Ca^{2+}-binding site I and helix II and is called the "S100-specific" or "pseudo-EF-hand". The C-terminal EF-hand comprises helix III, Ca^{2+}-binding site II – called "canonical" Ca^{2+}-binding motif, common to all EF-hand proteins, and helix IV. The N-terminal EF-hand is different from the canonical EF-hand motif and is characteristic for the S100 proteins [61]. Both EF-hand motifs are separated by a flexible hinge region, also known as a linker region. The binding of Ca^{2+} to site I induces changes in the backbone conformation, leading the protein to adopt a "Ca^{2+}-ready" state. This transition involves a rotation in helix III, resulting in a more open structure. This exposes a wide hydrophobic region encompassing residues from helices III and IV in the

C-terminal EF-hand, as well as the hinge region. This conformational change regulates the protein's activity by enabling the respective S100 protein to engage with a diverse array of target proteins, including receptors and other S100 members. [62]

The N-terminal EF-hands of some S100 proteins, such as calbindin, S100B, and S100A6, exhibit small conformational changes upon Ca^{2+}- binding. Helix II undergoes a small rearrangement in direction of the Ca^{2+}-binding site with practically no change in the interhelical angle of helix I and helix II (**Figure 1.10A**). The C-terminal canonical EF-hand represents the target interaction sites of the Ca^{2+} sensor S100 proteins. In the Ca^{2+}-free state, the helices III and helix IV flanking the EF-hand loop adopt an antiparallel conformation (**Figure 1.10B**). Upon calcium binding, the movement of helix III results in a perpendicular orientation of helices III and helix IV, changing the interhelical angle by approximately 90°. The Ca^{2+} induced conformational change opens the structure and exposes a hydrophobic cleft formed by residues of the hinge region, helix III, and the C-terminal loop region. [61]

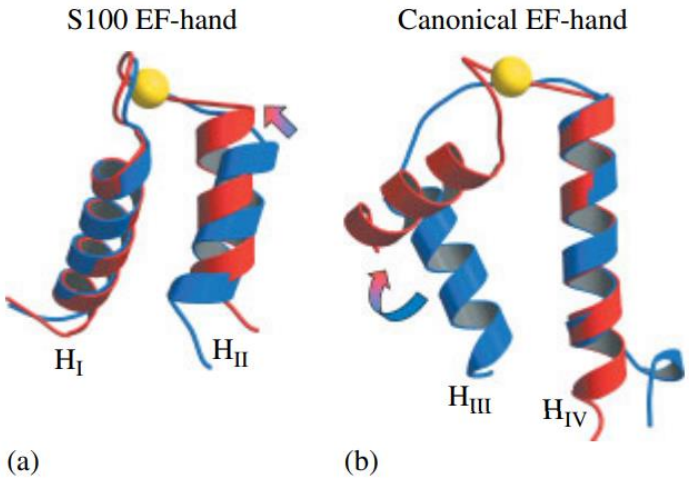

Figure 1.10 – Conformational changes in EF-hands of S100 proteins. The S100 specific EF-hand is depicted in **(a)**, the canonical EF-hand in **(b)**. The protein without bound Ca^{2+} is shown in blue, while the Ca^{2+}-loaded state is represented in red. Ca^{2+} ions are depicted as spheres coloured in yellow. [61]

S100 proteins are small, acidic proteins, with each protein being encoded by an individual gene. Among the 25 human S100 genes, 19 (known as group A S100 proteins) are located within chromosome 1q21. Other members such as S100A11P, S100B, S100G, S100P, and S100Z are distributed across distinct genomic regions. Every member of the S100 protein family has a similar molecular mass of 10–12 kDa, and they each share 25–65% similarity in their amino acid sequence [62].

While their distribution is not restricted to the nervous system, several members of the family are implicated in nervous system development, function, and disease [65]. S100 proteins can mediate inflammation and act as markers for inflammatory diseases. In fact, in neurodegenerative disorders, neuroinflammation is a common feature and is believed to contribute to disease progression [62]. Furthermore, some S100 proteins, particularly S100B, have been investigated as potential biomarkers for neurodegenerative disorders. Changes in their levels in biological fluids, such as cerebrospinal fluid, have been observed in various neurological conditions [66]. Furthermore, S100A9 [67], S100B [62] [66] and S100A6 [68] have been implicated in various neurodegenerative disorders, including Alzheimer's disease, Parkinson's disease, or multiple sclerosis. Monitoring S100 protein levels could potentially aid in the diagnosis and prognosis of these disorders. Building upon this foundation, the focus will be subsequently centred on the S100B protein characteristics and on its distinctive role within the context of neurodegenerative disorders.

1.3.1. S100B protein

The S100B protein, a member of the S100 family, stands as a key contributor to the intricate and complex network of cellular signalling and regulation. With its unique structure and functions, S100B has gained substantial attention in the field of neuroscience and neurology. This protein's multifaceted roles extend beyond its traditional function as a calcium-binding protein, encompassing a wide spectrum of activities that contribute to both normal physiological processes and disease-related conditions. S100B is the predominant S100 found in brain and makes up approximately 0.5% of all brain proteins. Under normal physiological states, S100B is expressed primarily in astroglial cells of the central nervous system, and is also found, although in a lesser extent, in neurons, microglia,, oligodendrocytes, neuronal progenitor cells, dendritic cells, kidney epithelial cells, and others. [60] [69] [70]

1.3.1.1. Structure and Function

The structural attributes of S100B parallel the fundamental framework delineated for S100 proteins earlier in this thesis. The S100B protein is characterized by its calcium-binding properties, which contribute to its functional versatility within the nervous system. It is composed of two EF-hand motifs, each consisting of a helix-loop-helix configuration. These EF-hand domains are responsible for binding calcium ions, which trigger conformational changes in the protein. S100B typically forms a dimeric structure, where two individual S100B protein subunits come together. This dimerization is often facilitated by interactions between the EF-hand motifs and the binding of calcium ions. The dimeric arrangement can influence S100B's ability to interact with various target proteins and mediate cellular responses. S100B exists primarily as a homodimer, composed of two monomers of approximately 11 kDa each, although stable and active tetrameric, hexameric, and octameric forms have also been reported [71].

The S100B monomer consists of four α-helices with a β-strand between both helices I and II, and III and IV, composing two helix-loop-helix EF-hand motifs connected by a linker or hinge region. The C-terminal canonical motif is made up of 12 amino acids involved in Ca^{2+}-ion binding, while the N-terminal EF-hand contains 14 amino acids, therefore being considered by some as a pseudo-EF-hand due to the extra two amino acids. Calcium binding in the N-terminal site I induces limited changes to the whole structure, while binding at the C-terminal site II induces a conformational change in helix III of up to 90°, leading to the exposure of the binding site for target proteins to access (**Figure 1.11**). [71]

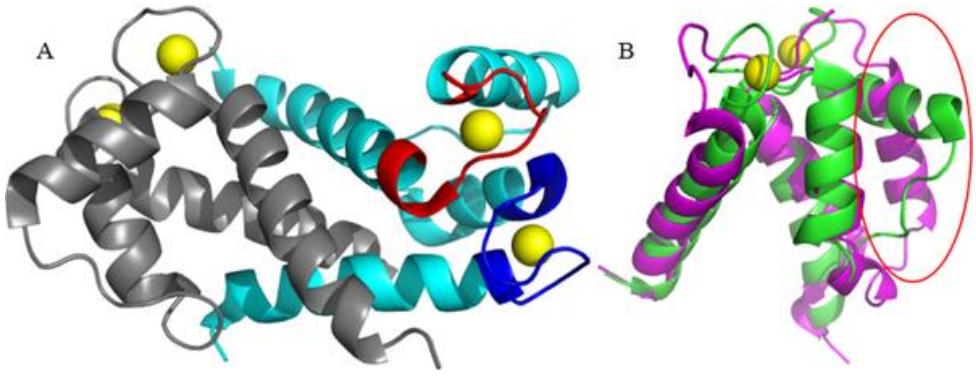

Figure 1.11 – Representation of S100B dimer and calcium-bound induced conformational changes. (A) S100B Ca^{2+}-bound dimer in ribbon representation (PDB entry 2h61). Monomers are coloured in grey and cyan. The C- terminal EF-hand motif of the cyan monomer is coloured in red, while the N-terminal pseudo-EF-hand motif is coloured blue. EF-hand bound Ca^{2+}− ions are shown as yellow spheres. **(B)** Alignment of apo-S100B (purple; PDB entry 1b4c) and Ca^{2+}− bound (green; PDB entry 2h61) S100B, shown as in A. Ca^{2+}− induces the rearrangement of helix III, highlighted with a red oval. (Adapted from [71]).

Like other calcium-binding proteins within the S100 family, S100B plays an important role in numerous intracellular processes by engaging with diverse molecules across distinct tissues and cell-types. S100B is involved in the regulation of calcium homeostasis, maintaining the balance of calcium ions within cells, as previously mentioned, but it also regulates cell proliferation and differentiation, cell survival, energy metabolism, enzyme activities, inflammation and migration/invasion trough interactions with various proteins (**Figure 1.12**). S100B has both intracellular and extracellular functions. It interacts with a variety of target proteins, including enzymes, cytoskeletal subunits, receptors, transcription factors, and nucleic acids [72] [73].

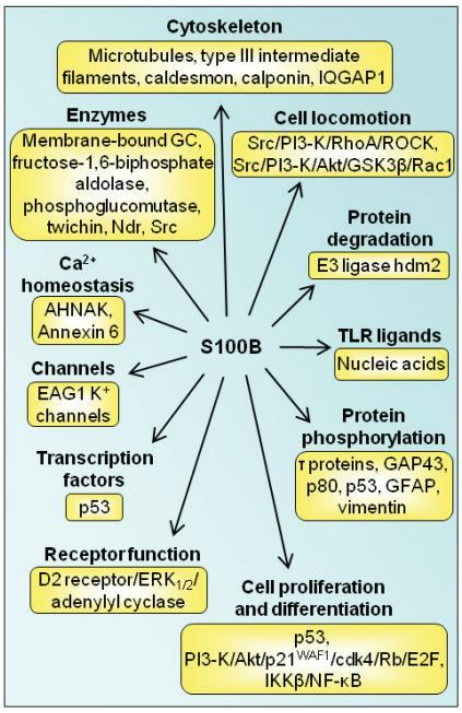

Figure 1.12 – Schematic representation of proposed intracellular effects induced by S100B. S100B establishes interactions with a multitude of intracellular proteins, modulating diverse cellular processes. It is involved regulation of protein phosphorylation, enzyme activities, the state of assembly of specific cytoskeleton components, the transcription factor p53, protein degradation, cell proliferation, migration and differentiation, dark adaptation of photoreceptors, the maintenance of Ca2+ equilibrium and the innate inflammatory response. [73]

As shown, S100B has a multifaceted role in orchestrating a spectrum of intracellular responses, depending on its interacting partner. Extracellularly, S100B functions are related mostly to signalling. Indeed, increasing evidence indicates that S100B exerts functional roles by acting as an intracellular regulator and an extracellular signalling molecule.

S100B is secreted by several cell types and released by damaged cells. Once released, S100B exerts regulatory effects on a relatively larger number of cell types in an autocrine, paracrine and, possibly, endocrine manner [74]. S100B secreted or released from astrocytes has different (trophic and toxic) effects on neurons, astrocytes and microglia depending on the concentration. Extracellularly, S100B interacts with adjacent different cell-types through the Receptor for Advanced Glycation Endproducts (RAGE), a transmembrane immunoglobulin-like receptor that binds to a diverse range of extracellular effectors, initiating a complex intracellular signalling cascade. This process has been implicated in various pathological conditions, including neuroinflammatory responses to neural injuries, as well as concomitantly resulting in an up-regulation of RAGE itself. Research has also reported that extracellular S100B can be re-uptaken by vesicles to astrocytes in a RAGE-dependent manner. Furthermore, studies

performed in cultured myoblasts (non-neuronal experimental system), S100B has been demonstrated to interact with the basic fibroblast growth factor (bFGF)/FGF receptor 1 (FGFR1) system. [66]

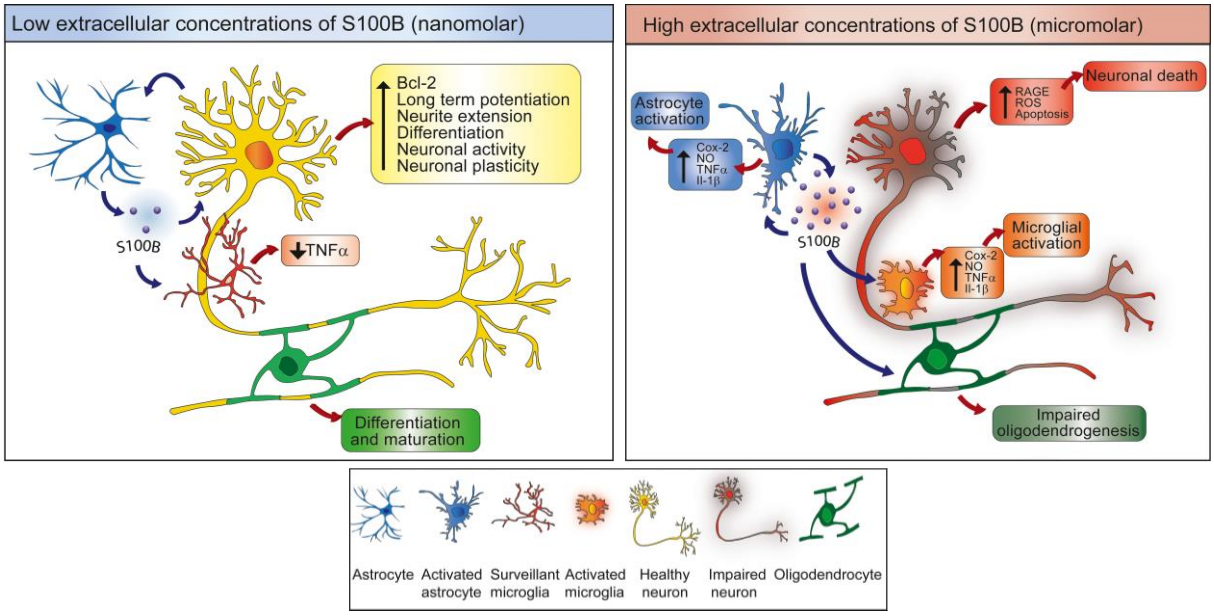

Figure 1.13 – Extracellular biological activity of S100B in the different cell types of the central nervous system, both at low and high concentrations. Low doses of S100B have been shown to protect neurons against apoptotic stimuli, enhance neurite outgrowth, among others, exhibiting a neurotrophic effect. At higher concentrations, S100B has been shown to have a neurotoxic effect, contributing to neuronal cell death, impaired oligodendrogenesis, astrocyte activation and microglia inflammation. (Adapted from [66])

At relatively high doses, S100B stimulates the mitogenic signalling of bFGF/FGFR1 in low-density myoblasts, upon the presence of bFGF. In this scenario, the S100B protein functions as a signal release from damaged muscles, contributing to the regeneration of skeletal muscle through the stimulation of either the promyogenic RAGE pathway or the bFGF/FGFR1 mitogenic pathway, depending on its own concentration, the absence or presence of bFGF, and myoblast density. The biological activity of extracellular S100B has been intricately associated with its concentration. Previous studies, subsequently corroborated and confirmed, indicated that S100B has a neurotrophic effect at low (nanomolar) concentrations, which are considered physiological. Effects such as promotion of neurite extension, modulation of long-term potentiation, protection of neuron viability, mitigation of neurotoxicant insults, increased scavenger activity of reactive oxygen species (ROS) have been reported. Nanomolar concentrations of S100B induce the generation of low amounts of signalling oxygen radicals, culminating in the activation of the anti-apoptotic factor Bcl-2. [66]

On the contrary, S100B has demonstrated toxic and pro-inflammatory effects at elevated concentrations, in the micromolar range. Prolonged activation of RAGE by micromolar concentrations of S100B results in increased amounts of oxygen radical, potentially leading to mitochondrial dysfunction and induction of apoptosis. Reports indicate that micromolar concentrations of S100B up-regulate the expression of inducible nitric oxide synthase (iNOS), triggering the release of nitric oxide (NO) and NO-dependent death of neurons and glial cells. Additionally, these concentrations are linked to exacerbate glutamate-induced neuronal death, to intensify activation states in astrocytes, to increase cyclooxygenase-2 (COX-2) expression in microglia, to elevate reactive oxygen species (ROS) production in neurons, to cause disturbances in lipid homeostasis and cell cycle arrest, and to impair oligodendrogenesis [66]. All these functions are represented in **Figure 1.13**.

RAGE has been identified as an S100B receptor, transducing S100B effects on a variety of cell types with different outcomes (i.e. beneficial or detrimental, pro-proliferative or pro-differentiative) depending on the concentration attained by the protein and the cell type. S100B activates a RAGE-dependent autocrine loop in astrocytes, turning them into a pro-inflammatory/neurodegenerative phenotype. The different effects exerted by S100B at low and high concentrations on nervous cells likely depend on the level of RAGE expression, different intensities of RAGE activation, the duration of RAGE stimulation and/or different extents of S100B-induced upregulation of RAGE expression in neurons, astrocytes and microglia. Environmental parameters such as the Ca^{2+} and/or ROS concentration and the S100B clearance rate also might influence S100B/RAGE final effects [74]. Thus, S100B, in common with other proteins of the S100 family, may be joined to danger/damage-associated molecular pattern (DAMP) molecules, or alarmins, which are released in the endogenous microenvironment to trigger tissue reaction to damage. Interestingly, some characteristics of S100B, such as its non-canonical secretion modality that bypasses the classical Golgi route, its interaction with RAGE and its ability to stimulate microglial migration, are shared with DAMPs. Since DAMPs can be released in the extracellular space in the early stages of diseases and their levels often correlate with the disease progression, they may be used as biomarkers for the appropriate diagnosis and prognosis of different pathologies [66]. Therefore, S100B acts as a biomarker for conditions such as traumatic brain injury (TBI) since it's released at high concentrations during neural injury [66], as recognized and approved by the Scandinavian Neurotrauma Committee [75].

1.3.1.2. S100B in neurological diseases

Alzheimer's Disease

S100B expression has been reported to be enhanced in several tumours from cell-types normally expressing the protein, as well as in the aging brain [74]. Additionally, astrocytic S100B has been shown to be upregulated in brain tissues of AD patients, and its aberrant levels, acting as a neurotrophic factor, have been considered a potential contributor to the increased concentration of aggregates of overgrown neurites in the neuritic plaques. Furthermore, the detection of S100B in biological fluids has gained recognition as a dependable biomarker for AD. Notably, S100B levels in cerebrospinal fluid, in conjunction with other AD biomarkers like amyloid-β and phosphorylated-τ, have recently revealed distinct correlations with increased grey matter volumes and heightened glucose metabolism in crucial Alzheimer-related regions. [72]

Alzheimer's disease (AD) is a progressive neurodegenerative disorder and the most common cause of dementia. It is characterized by cognitive decline, memory loss, and impaired daily functioning. The two hallmark pathological features of Alzheimer's disease are the accumulation of abnormal proteins in the brain: tau protein tangles and amyloid-beta (Aβ42) plaques. [76]

Interestingly, S100B's chaperone activity has been recognized in the context of neurodegenerative diseases and is thought to play a protective role in neuronal health. S100B has been found to inhibit Aβ42 aggregation in a Ca^{2+}-dependent manner, a crucial step in the formation of toxic amyloid plaques, which is a pathophysiological hallmark of Alzheimer's disease [77-81]. This inhibitory effect is even more pronounced when S100B interacts with Aβ42 in the presence of heparin. S100B acts as a holdase-type chaperone, preventing not only Aβ42 aggregation but also the initiation or seeding process necessary for amyloid propagation [80] [79]. Additionally, a recent study proved that at nanomolar concentrations and in the presence of calcium, S100B reduces the toxicity of Aβ42, and delays Aβ42 aggregation by interacting with both Aβ42 monomers and fibrils in SH-SY5Y cells (**Figure 1.14**) [81] [78]. Furthermore, research has explored the potential of targeting S100B with specific peptides encoding aggregation-prone sequences. These peptides can modulate S100B's function, preventing Aβ42 aggregation, especially under conditions that promote protein misfolding and aggregation (destabilizing conditions) [82], which are central to Alzheimer's disease pathogenesis.

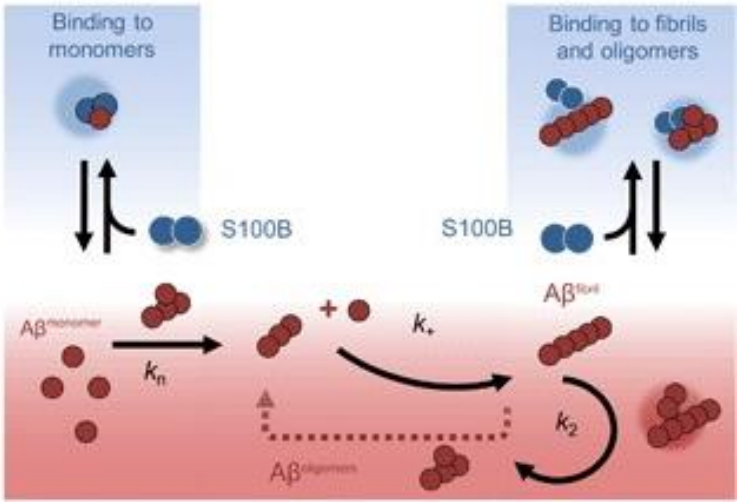

Figure 1.14 – Inhibitory effect of the S100B chaperone on Aβ42 peptide amyloid aggregation. Extracellular S100B acts as a chaperone in the presence of calcium, inhibiting the formation of aggregates by interaction with Aβ42 monomers, oligomers, and fibrils. [81]

Tauopathies

Moreover, S100B has been found to inhibit the aggregation of tau protein. Both full-length tau and the microtubule binding domain aggregation have been delayed upon the presence of S100B, in a Ca^{2+}-dependent manner at sub-stoichiometric ratios and demonstrated to have a protective role in preventing tau inclusion formation induced by toxic tau oligomers (**Figure 1.15**). [80] [79]

Furthermore, a recently published study demonstrated that S100B plays a role in modulating tau liquid-liquid phase separation (LLPS), induced by crowding agents and metal ions. In the absence of calcium (apo S100B), its effect on PEG-induced tau demixing is minimal. However, when bound to calcium ions (Ca^{2+}-bound S100B), it effectively prevents tau demixing. This study has also revealed that S100B incorporates into tau liquid droplets, indicating its crucial stabilizing and chaperoning role, which contributes to minimizing the formation of toxic tau aggregates. [83]

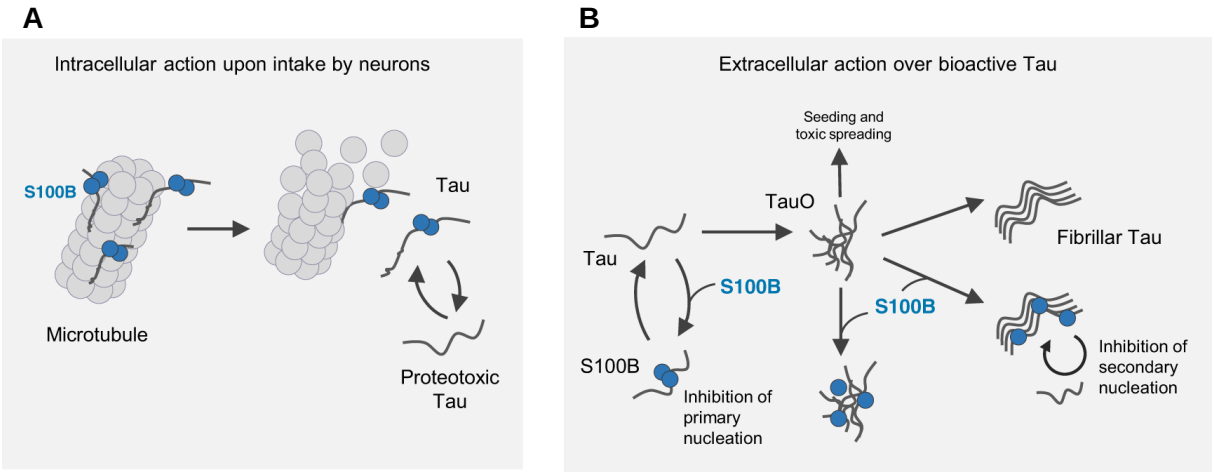

Figure 1.15 – Proposed representation of the model for the intra- and extracellular action of the S100B chaperone on tau. (A) Intracellular interaction of S100B with tau, with S100B operating as a microtubule-regulating factor, as a protective holdase, that mitigates protetotoxic tau that has been released from the microtubules after their destabilization. **(B)** In the extracellular environment, levels of S100B increase in the early stages of Alzheimer's disease and are prominently secreted by nearby astrocytes in the presence of tau-positive inclusions. In this extracellular context, S100B acts as a chaperone, impeding tau aggregation by engaging with released proteotoxic tau. This interaction contributes to reducing the initiation of pathological seeding events and limits the spread of tau pathology from one cell to another. [79]

Parkinson's Disease

S100B has been found to have an emerging role in the pathogenesis of Parkinson's disease (PD), a neurodegenerative disorder characterized by motor symptoms such as tremors and rigidity. S100B has been reported to be overexpressed in astrocytes in the brain tissue of PD patients, as well as in animal PD models obtained using 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP), a neurotoxin shown to produce neurological and pathological changes comparable to those observed in Parkinson's disease [74]. Additionally, research has revealed that mice with elevated S100B levels are more susceptible to developing Parkinsonian symptoms. Furthermore, exposure to S100B in midbrain cultures has been observed to specifically affect the activity of dopaminergic neurons expressing tyrosine hydroxylase (TH^+). Interestingly, elevated S100B levels overnight have been linked to increased severity of Parkinson's disease (PD) symptoms and disruptions in sleep patterns. Additionally, specific S100B gene variations appear to be associated with the age at which PD symptoms first appear. There is also ongoing exploration of the potential use of S100B as a biomarker in biological fluids to monitor the progression of PD based on findings from studies involving PD patients. Furthermore, in experimental animal models of PD, the use of an inhibitor targeting S100B activity, PTM, resulted in a significant improvement in motor performance. In studies involving S100B knockout mice treated with MPTP to induce PD, the absence of S100B expression was accompanied by improvements in various pathological parameters, including reduced loss of dopaminergic neurons, decreased microgliosis, and decreased expression of tumor necrosis factor alpha [72].

Amyotrophic lateral sclerosis

Amyotrophic lateral sclerosis (ALS), is a rare neurodegenerative disease that results in the progressive loss of motor neurons that control voluntary muscles, ultimately leading to paralysis [84]. Elevated levels of the S100B protein have been observed in neural tissues, particularly in astrocytes, within ALS patients. Additionally, in a rodent model of ALS, there have been reports of increased expression of S100B, its receptor RAGE, and the High Mobility Group Box 1 protein—another DAMP molecule binding to RAGE—in the lumbar spinal cord [72]. Thus, a potential role for these molecules in the progression of ALS has been further proposed [85].

Multiple Sclerosis

Earlier data correlating S100B to Multiple Sclerosis (MS) were obtained in the late 1970s, when high levels of the protein were detected in the cerebrospinal fluid (CSF) of MS patients during the acute phase, while levels were normal in the remission phase [72]. Multiple sclerosis is a chronic neurodegenerative and autoimmune disease of the central nervous system characterized by inflammation, demyelination, gliosis, and neuronal loss [86]. It is characterized by progressive axonal damage, leading to the gradual loss of motor neurons and resulting in severe muscle weakness and paralysis [87]. Furthermore, elevated levels of the S100B protein were observed in the serum of individuals with multiple sclerosis at disease onset. This presence of S100B protein in the biological fluids of MS patients correlated with characteristics seen in the nervous tissue of individuals with MS. In postmortem tissues from these patients, increased S100B expression was detected in both active demyelinating and chronic active MS plaques, and, in addition, there was an overexpression of RAGE in active demyelinating lesions. These data proposing the involvement of the S100B protein in MS were confirmed in experimental models of the disease [72].

Due to its involvement in a wide range of neurodegenerative diseases, as described, including AD, PD, ALS and MS, S100B has attracted significant attention in the field of neurology. Its influence goes beyond its overexpression in the brain in all these conditions, but also in the fact that it has been observed to directly influence the aggregation dynamics of proteins, a fundamental process underlying the pathogenesis of many of these conditions. Given this, it becomes plausible that S100B could exert a similar influence on huntingtin aggregation, raising the possibility that it may also contribute to the pathology of HD, potentially influencing the disease's onset and progression. Exploring the relationship between S100B and huntingtin in the context of HD could yield valuable insights into the complex mechanisms at play in this devastating condition.

1.4. Objectives

Huntingtin aggregation is a defining characteristic of Huntington's Disease. The S100B protein, primarily synthesized and expressed by astrocytes, has recently been discovered to possess a novel chaperone function that inhibits amyloid aggregation and alleviates associated toxicity. Interestingly, at the biologic concentrations found in a healthy brain, S100B is proposed to exert a protective role, functioning as a molecular chaperone capable of suppressing the aggregation of amyloidogenic proteins, becoming a potential therapeutic target to mitigate such conditions [81]. Thus, S100B has emerged as a novel player in maintaining brain proteostasis with possible implications in other neurodegenerative diseases, such Huntington's Disease.

The objective of this study was to investigate regulatory interactions between huntingtin and S100B proteins with impact on HD pathophysiology, combining biochemical, biophysical and cell biology methods. The two following approaches were pursued:

Task 1 – Cellular and interaction assays:

A) Live-cell assays to investigate the interactions between S100 proteins and Htt using a dual luminescence-based co-immunoprecipitation assay, for systematic detection of protein-protein interactions in mammalian cells (LuTHy).
B) Evaluate effects of endogenous and exogenous S100B (S100B-Cerulean and S100B-Myc, respectively) on live cells expressing wild-type and mutant huntingtin under punctae/aggresomes forming conditions using combining biochemical and bimolecular fluorescence complementation assays (BiFC).

Task 2 – Htt aggregation studies:

A) Expression of recombinant human S100B-Myc proteins from bacterial cultures and purification using chromatographic methods.
B) Expression of wild-type and mutant huntingtin (HttExon1-Q23 and HttExon1-Q43, respectively).
C) Characterization of S100B-Myc using biophysical methods – Circular Dichroism and ANS fluorescence spectroscopies.
D) *In vitro* assays to investigate the influence of S100 proteins on Htt aggregation using a fluorescence resonance energy transfer (FRET)-based Htt aggregate seeding (FRASE).

II. Materials and Methods

2.1. HttExon1-Q23 and HttExon1-Q43 expression

The plasmids pTWIN1-His6-Ssp-HttExon1-23Q-A2C-A60C and pTWIN1-His6-Ssp-HttExon1-43Q were obtained from Addgene incorporated in NEB cells. In order to purify the plasmids from the NEB cells, a kit was used (NZYMiniprep, NZYTech), followed by the determination of the DNA concentration spectrophotometrically (Abs_{260nm}). The expression and purification of HttExon1-Q23 and HttExon1-Q43 were performed based on the previously established protocol of Lashuel *et al* with some modifications [88].

Escherichia coli (DE3) BL-21 competent cells were transformed with 1 µL of each plasmid in two distinct sterile culture tubes. For transformation, heat shock was performed (incubation on ice for 30 minutes, followed by 45 seconds at 42 ºC and 2 minutes on ice). Next, LB was added, and the cells were incubated at 250 rpm for 1 hour at 37 ºC. The transformed cells were grown in plaques with Luria-Bertani Agar medium (LA) with 100 µg/mL of ampicillin overnight at 37 ºC. The next day, a single isolated colony was grown in 25 mL of LB in the presence of 100 µg/mL of ampicillin and incubated at 37 ºC, 200 rpm until an optic density at 600 nm (OD_{600nm}) between 0.6 and 0.8 was reached. At this point, glycerol stocks were made, with the freezing of the samples on liquid N_2 and the samples were stored at -80 ºC.

For the expression, from the glycerol stocks, cells were inoculated in 250 mL of LB containing 100 µg/mL of ampicillin and were incubated overnight at 37 ºC, 180 rpm. LB containing 100 µg/mL of ampicillin was added to the overnight culture to obtain an OD_{600nm} = 0.05. The cell culture was then incubated at 37 ºC, 180 rpm shaking until an OD_{600nm} = 0.1 was reached, and after the temperature was set to 14 ºC with the same shaking. When OD_{600nm} = 0.3 – 0.4, expression was induced with 0.4 mM of isopropyl β-d-1-thiogalactopyranoside (IPTG) and was left incubating overnight (14 ºC, 180 rpm). On the following day, the cells were harvested by centrifugation (4 ºC, 8000 rpm, 10 min, JA-14 rotor), snap-frozen in liquid N_2 and stored at -20 ºC.

2.2. S100B-Myc dimer purification

Human recombinant S100B-Myc protein was expressed and purified following the previously established protocols [89] [90]. S100B-Myc had previously been overexpressed in BL21 (DE3) competent *E. coli* (Lucigen) cells with 1 µL of plasmid pGEMEX-S100B-Myc (Origene®) encoding the human wild-type S100B-Myc protein and the pellet stored at -20ºC.

For the transformation, cells were incubated on ice for 30 minutes, followed by heat shock for 45 seconds at 42 ºC, and 2 minutes on ice. Then, LB was added, and the cells were incubated at 250 rpm for 1 hour, 37 ºC. The transformed cells were plated in plaques with LA medium in the presence of 100 µg/mL of ampicillin for overnight growth at 37 ºC. On the following day, a single colony was inoculated in 100 mL of LB containing 100 µg/mL of ampicillin at 37 ºC and 150 rpm overnight. The next day, 10 mL of culture were added to 500 mL of LB in the presence of 100 µg/mL of ampicillin and incubated at 37 ºC, 150 rpm. When OD_{600nm} = 0.6 – 1, protein expression was induced with 1 mM of IPTG, followed by incubation at 37 ºC, 150 rpm, ON. The cells were harvested by centrifugation (Beckman Coulter J2-21, JA-14 rotor) at 8000 rpm for 10 minutes, at 4 ºC and stored at -20 ºC.

For the purification, the cell pellet was thawed, and cells were resuspended with resuspension buffer (50 mM Tris-HCl, pH 7.6, 5mM $MgCl_2$) complemented with DNAse (Panreac) and 0.5mM of phenylmethanesulfonyl fluoride (PMS) (Roth) and protease inhibitor cocktail (1 tablet, Roche). Then, cells were disrupted using a high-pressure French Press/Carver at 2000 psi and centrifuged (Beckman L8-70M ultracentrifuge, 45Ti rotor) at 42000 rpm, 4 ºC for 1 hour to remove insoluble material. The resulting supernatant was diluted 2-fold with phenylsepharose buffer A (50 mM Tris-HCl, pH 7.6, 5 mM $CaCl_2$) and was left at 4 ºC for 2h with slow agitation. The following day, a 20mL PhenylSepharose6 column (HiPrep Phenyl Fast-Flow, Cytiva) was equilibrated with the same buffer in order to perform a hydrophobic interaction chromatography using a FPLC ÄKTA purifier UPC 10 system (GE Healthcare/Cytiva) and column loading was done using a superloop. The protein sample was filtered (syringe filter 0.45 mm) and injected into the column at a flow rate of 2mL/min, allowing the complete elution of the unbound fraction. The elution of S100B-Myc was performed after Abs280nm reached basal levels, through switch from buffer A to phenylsepharose buffer B (50 mM Tris-HCl pH 7.6, 10 mM EDTA) at a flow rate of 5 mL/min. The fractions corresponding to the UV peaks, were collected for further analysis by SDS-PAGE (8% Tris-Tricine), and the peaks corresponding to S100B-Myc were pooled into a single fraction and concentrated to 10 mL in a 3 kDa cut-off Amicon® centrifuge filter (Millipore, Merk). A size exclusion chromatography was performed and a Superdex 75 16/600 column (Cytiva) was equilibrated overnight with a buffer containing 20mM Tris-HCl, pH=7.6 and 150 mM NaCl at a flow rate of 0.5 mL/min. Using a superloop, the concentrated volume containing S100B-Myc was injected into the column and eluted at 1 mL/min. The fractions corresponding to the UV peaks were collected for analysis by SDS-PAGE (8% Tris-Tricine) and fractions corresponding to the S100B-Myc dimer were pooled into a single fraction, further concentrated in the same cut-off Amicon® centrifuge filter. Concentration was calculated spectrophotometrically by Abs_{280nm} (ε_{280nm} S100B-Myc dimer) = 5960 $M^{-1}cm^{-1}$) and the S100B-Myc was aliquoted in 2 mL *eppendorfs*, flash-frozen and stored at -20 ºC.

Apo S100B-Myc dimer was obtained through incubation of the sample with a 300-fold excess of DTT and 0.5 mM EDTA (37 ºC, 2 hours) and elution in a Superdex 75 Tricorn high-performance column (Cytiva) with 50 mM Tris-HCl, pH 7.4 and 150 mM NaCl chelexed buffer. Prior to its injection on the column, sample was centrifuged at 14500 rpm for 5 minutes (benchtop microcentrifuge) to remove suspension particles. The protein was eluted at a 1 mL/min flow rate, the peak corresponding to the dimer peak was collected and concentrated in a 3 kDa cut-off Amicon® centrifuge filter, flash-frozen and stored at - 20 ºC. Sample concentration was calculated spectrophotometrically by Abs_{280nm} (ε_{280nm} S100B-Myc dimer) = 5960 $M^{-1}cm^{-1}$).

2.3. Biophysical characterization of S100B-Myc

2.3.1. Circular dichroism assays

Circular dichroism (CD) measurements were performed on a Jasco J-1500 spectropolarimeter equipped with a Peltier-controlled thermostated cell support at 25ºC. Samples were prepared by diluting S100B-Myc to a final concentration of 5µM (0.1 mg/mL) in 20 mM Tris pH 7.6 150 mM NaCl and 50 mM Tris-HCl pH=7.4, chelexed, for the holo and apo forms of the protein respectively. The thermal stability was assessed by thermal unfolding with a linear temperature increase of 1ºC/min from 20 to 98 ºC, following ellipticity variation at 200 nm and 222 nm. Far-UV CD were recorded between 195/205 nm and 260 nm using a 1 mm pathlength quartz cuvette (Hellma Analytics) and 10 accumulations. Background buffer baselines were subtracted from each spectrum and data represented in a plot using the Origin software.

2.3.2. ANS assays

ANS (8-Anilino-1-naphtalenesulfonic acid) measurements were performed on a Jasco FP8200 spectrofluorometer at 25ºC. Samples were prepared by diluting S100B-Myc in the apo form to a final concentration of 10µM with two-fold of ANS in 50 mM HEPES pH 7.4 for 2h at 25ºC. Calcium was added to a final concentration of 1mM and 2 mM and incubated for additional 5 minutes at 25ºC. ANS emission spectra was recorded between 375 nm and 650 nm using an excitation wavelength of 370 nm, 5 nm bandwidth, medium sensitivity and 5 scans of average accumulation. Data was represented in a plot using the Origin software.

2.4. FRASE assay

These experiments were performed at the MDC, Berlin at the Proteomics and Molecular Mechanisms of Neurodegenerative Diseases laboratory, under the guidance of Dr. Erich Wanker.

The FRASE assay was performed based on the previously established protocol by Ast. *et al* [91]. Purified GST-Ex1Q48-CyPet and GST-Ex1Q48-YPet proteins were thawed on ice. Then, the proteins were transferred into ultracentrifuge tubes and centrifuged for 40 minutes at 55000 rpm, 4°C (Centrifuge: Optima TLX Ultracentrifuge, TLA55 rotor). After centrifugation, 70% of the centrifuged volume of GST-HttExon1-Q48CyPet/YPet was collected, and the pellet discarded. Protein concentration of the supernatant was determined using Nanodrop (Thermofisher) and were entered into the protocol in order to determine the volume of protein needed to prepare the mixes (background, donor, acceptor and FRET Master mixes) in LoBind Eppendorf tubes. The mixes were then vortexed on a low setting to ensure equal distribution of all components. The FRET Master mix was prepared containing: buffer A (50 mM Tris, 150 mM NaCl, 1 mM EDTA, 5 % Glycerol, pH 7.4); 50µM $CaCl_2$; distilled H_2O; DTT 100mM (Serva), PreScission protease (500 U, GE Healthcare) and 5µM of GST-Ex1Q48-CyPet and GST-Ex1Q48-YPet proteins. The FRET Master mix was distributed onto the samples. All samples were vortexed collectively within the rack and are further loaded onto a 384-well-plate (Thermo Scientific™ Nunc™ 384-Well Polystyrene Black Microplates) - 30µl/well (Background, Donor and Acceptor mixes in singlets/all FRET samples in triplicates). Fluorescence intensities were measured using a microplate reader (TECAN infinite 200) at 25 °C for up to 80 h. The CyPet donor fluorescence was measured at excitation (Ex): 435 nm/emission (Em): 475 nm; YPet acceptor fluorescence at Ex: 500 nm/Em: 530 nm; the FRET channel (DA) was recorded at Ex: 435 nm/Em: 530 nm. For data processing, raw signals were processed by subtracting the background fluorescence (blank) in all channels. Signals in the FRET channel were corrected for donor bleed-through (cD) and acceptor cross excitation (cA) using donor- and acceptor-only samples to obtain sensitized emission (donor and acceptor mixes). Finally, sensitized emission was normalized to the acceptor signals. In brief, the FRET efficiency E (in %) was calculated as follows:

$$E = \frac{(DA - cD \times DD - cA \times DA)}{AA}$$ with DD = donor channel signal and AA = acceptor channel signal.

The FRET efficiency was further normalized and represented in a plot in Origin software.

2.5. LuTHy assay

These experiments were performed at the MDC, Berlin at the Proteomics and Molecular Mechanisms of Neurodegenerative Diseases laboratory, under the guidance of Dr. Erich Wanker.

The LuTHy assay was performed based on the previously established protocol by Trepte *et al* [92].

Plasmid construction For LuTHy experiments, all cDNAs were shuttled from the entry vectors into the LuTHy destination vectors (pcDNA3.1) using the LR clonase technology following the instructions of the manufacturer (Invitrogen).

Cell culture and transfection The human embryonic kidney cell line 293 (HEK293) was grown in high glucose (4.5 g/l) DMEM (Gibco®, ThermoFisher). Medium was supplemented with 10% heat inactivated fetal bovine serum (Gibco®, ThermoFisher), 1% Pen/Strep, and cells were grown at 37°C, 5% CO2. Cell subculturing was done every 3–4 days and transfection was carried out using linear polyethylenimine (25 kDa, Polysciences) through the reverse transfection method, following the instructions provided by the manufacturer. For LuTHy transfections, cells were seeded in phenol red-free, high-glucose DMEM media (Gibco®, ThermoFisher) supplemented with 10% heat inactivated fetal bovine serum. Transfections were executed with 200 ng of total DNA per well in a 96-well plate. If expression plasmid concentration was below 200 ng/well, pcDNA3.1(+) was used as carrier DNA to fill the total DNA amount to 200 ng.

LuThy assay HEK293 cells were reversely transfected in white 96 well microtiter plates at a density of 4.5x104 cells per well. Plasmids encoding donor and acceptor proteins were transfected at a 1:10 to 1:20 ratio, with 5–10 ng DNA for the donor and 100 ng for the acceptor. After 48h of the transfection, mCitrine fluorescence was assessed in intact cells, employing excitation/emission wavelengths of 500 nm and 530 nm, respectively. Coelenterazine-h (NanoLight, 301 or pjk, 102182) was added to the cell medium to a final concentration of 5 µM. Cells were incubated for an additional 15 minutes, and the total luminescence, along with the luminescence within short-wavelength (short-WL) and long-wavelength (long-WL) ranges, was measured. Fluorescence and luminescence were measured using the BLUE1 filter (370–480 nm) and GREEN1 filter (520–570 nm), which were utilized with an integration time set at 1,000 ms using Tecan's Infinite® microplate readers M200, M1000, or M1000Pro. Following luminescence measurements in intact cells, the luminescence-based co-precipitation (LuC) was carried out. Cells were lysed in 50–100 µl of HEPES-phospo-lysis buffer (50 mM HEPES, 150 mM NaCl, 10% glycerol, 1% NP-40, 0.5% deoxycholate, 20 mM NaF, 1.5 mM MgCl2, 1 mM EDTA, 1 mM DTT, 1 U Benzonase, Roche's EDTA-free protease inhibitor cocktail, 1 mM PMSF, 25 mM glycerol-2-phosphate, 1 mM sodium orthovanadate, 2 mM sodium pyrophosphate) for 30 minutes at 4°C. The generation of PA-mCit- and NL-tagged fusion proteins was monitored by assessing fluorescence ($mCit_{IN}$) and luciferase activity (NL_{IN}) within crude cell lysates using white, small-volume 384-well microtiter plates (Greiner, 784074). Coelenterazine-h was added (5 µL of coelenterazine-h) to 5 µl of cell lysates to achieve a final concentration of 10 µM, and luminescence was measured with a microplate reader with 100 ms integration time as previously described. Small-volume 384-well microtiter plates (Greiner, 784074) were prepared a day prior by coating them using sheep gamma globulin (Jackson ImmunoResearch, 013-000-002) antibody in carbonate buffer (70 mM NaHCO3, 30 mM Na2CO3, pH 9.6) for 3 h at room temperature. The plates were blocked with 1% BSA in carbonate buffer before incubating with rabbit anti-sheep IgGs in carbonate buffer (Jackson ImmunoResearch, 313-005-003) overnight at 4°C. All wells were equilibrated with lysis buffer directly before usage. Note that IgG-coated plates should not be stored for more than 24 hours. From the cell lysate, 15 µL of it was incubated for 3 h at 4°C in the IgG-coated 384-well plates. Subsequently, all wells were washed three times using

lysis buffer, and mCitrine fluorescence ($mCit_{OUT}$) was measured as described earlier. Finally, 15 µl of PBS buffer containing 10 µM coelenterazine-h is added to each well and the luminescence activity (NL_{OUT}) is measured as described above.

2.6. BiFC assay

These experiments were performed in Ciências/ULisboa, in collaboration and with guidance from Dr. Federico Herrera and at the BioISI microscopy facility.

Mammalian cell culture HeLa cells were maintained in DMEM (Thermo Fisher Scientific), supplemented with 10% FBS, 2mM of glutamine stable 100x (Biowest) and maintained at 37°C in a 5% CO2 atmosphere. For microscopy, 30.000 cells were seeded in 35 mm glass-bottom dishes (ibiTreat), and for protein extraction 800.000 cells were seeded in 60 mm dishes. Cells were transfected with the different combinations of plasmids using jetPRIME reagent (Polyplus) 16-24 hours after seeding, according to the manufacturer's instructions. Cells were transfected with 0.5 µg of each plasmid of wild-type HttExon1 fragments (19Q/25Q) and with 1 µg of each plasmid of mutant HttExon1 fragments (97Q/97Q).

Constructs The Htt constructs used, HttExon1-Q19-VN/HttExon1-Q25-VC and mutant HttExon1-Q97-VN/HttExon1-Q97-VC, were described elsewhere [93], as well as the S100B-Cerulean construct [79].

Microscopy Twenty-four hours after transfection, cells were observed by live imaging and pictures of fluorescent cells were taken with a Leica DMI6000 B widefield fluorescence microscope. Images were analysed using ImageJ free software (http://rsbweb.nih.gov/ij/).

Protein Extraction and Western Blotting Twenty-four hours after transfection, proteins were extracted by scrapping cells directly from the plates into lysis buffer. For denaturalizing conditions, the lysis buffer was 1% NP-40 (v/v), 150 mM NaCl, 50 mM Tris pH 7.4 and a protease inhibitor cocktail tablet (Roche) (denature lysis buffer). The native lysis buffer used was 50 mM Tris-HCl pH 7.4, 173 mM NaCl, 5 mM EDTA and a protease inhibitor cocktail tablet (Roche). Cells were then sonicated (12 pulses) and centrifuged at 10,000 x g for 10 min at 4 ºC. Supernatant was collected (soluble fraction), and the pellet was resuspended in denaturing lysis buffer, centrifuged and the resultant supernatant corresponded to the insoluble fraction of the cell lysate. Protein concentration was quantified by the Bradford assay (Thermofisher Scientific, Waltham, USA). Thirty µg of protein were loaded and run on 10% and 15% acrylamide SDS-PAGE for denaturing conditions. Proteins were transferred to nitrocellulose of PVDF membranes for blotting with antibodies against Huntingtin-P (phospho-N17) (1:1000, generated by Truant, R. [94], McGill University) and Huntingtin (1:1000, MAB5374, Sigma-Aldrich), GAPDH (1:1000, 6C5, sc-32233, Santa Cruz Biotechnology), Actin (1:1000, 2Q1055, sc-58673, Sigma-Aldrich), S100B β-chain (1:2000, S2532, Sigma-Aldrich) and Myc tag (1:250, M5546, Sigma-Aldrich) as indicate in each case. PVDF membranes were stripped (stringent stripping, ThermoFisher) and reprobed for the controls. Immunoblots were developed with enhanced chemiluminescence reagents (ECL) (GE Healthcare) and with further exposition in Amersham Imager 680 RGB.

III. Results and discussion

3.1. Htt-Exon1-Q23 and Htt-Exon1-Q43 expression

The nature of the N-terminal fragment exon 1 of huntingtin is highly prone to aggregation, particularly when it comprises an expanded polyQ stretch. As a result, producing HttExon1 using conventional recombinant expression systems in bacteria and mammalian cells faces substantial difficulties. Due to these challenges, most *in vitro* studies examining the structure and aggregation behaviour of Httex1 have taken alternative approaches, such as engineered fusion constructs where either the polyQ domain or the HttExon1 protein itself is merged with larger solubilizing protein tags. The enzymatic cleavage of HttExon1 fusion proteins frequently results in the incorporation of extra amino acids at either the starting (N-terminus) or ending (C-terminus) point of the protein. This change has the potential to modify the biophysical and biochemical properties of HttExon1, due to the important role the N17 and PRD domains display in regulating the structural, conformational and aggregation properties of the protein. To overcome these difficulties, an intein-based strategy was used to produce native tag-free HttExon1 [88]. An intein (intervening protein) is a segment of a protein that has the unique ability to self-catalyse its own removal from a larger protein precursor and ligates the flanking extein (external protein) sequences through the formation of a new peptide bond. This self-splicing process occurs at the peptide level without the requirement of external enzymes [95]. This intein-based strategy eliminates the need for proteolytic enzymes and has been successfully used to produce Httex1 with different polyQ lengths. The constructs follow the design: His6-Ssp-HttExon1-Q_N (N corresponds to the number of glutamines comprising the polyQ stretch, in this case 23 or 43, with molecular weights of 28.4 kDa and 30.9 kDa, respectively) After *E.coli* cell lysis, the soluble fraction was obtained, and the corresponding pellet was resuspended to obtain the insoluble fraction containing inclusion bodies with the overexpressed protein.

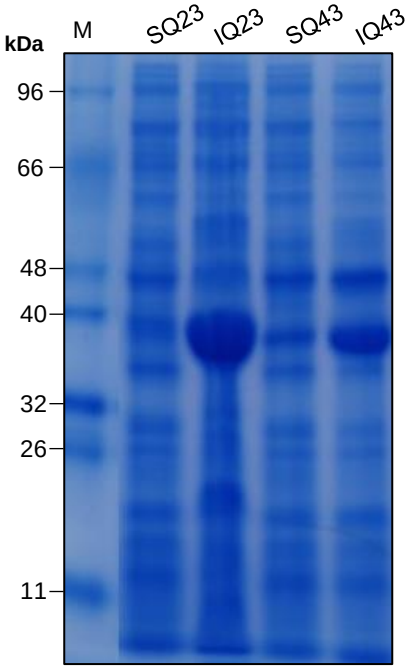

Figure 3.1 – SDS-PAGE 8% tris-tricine of soluble and insoluble fractions from E.coli extract overexpressing Htt-Exon1-Q23 and Htt-Exon1-Q43 proteins (4µg loaded in each well). MW: molecular weight markers (LMW II marker, NZYTech). SQ23- soluble fraction with Htt-Exon1Q23 overexpressed; IQ23 – insoluble fraction with HttExon1-Q23 overexpressed; SQ43- soluble fraction with Htt-Exon1Q43 overexpressed; IQ43 – insoluble fraction with HttExon1-Q23 overexpressed.

In **Figure 3.1**, the SDS-page electrophoresis gel is represented, in which the soluble and insoluble fractions of HttExon1-Q23 and HttExon1-Q43 were loaded. It is possible to observe that the lanes

corresponding to the insoluble fraction have an overexpressed protein relative to the soluble fractions, suggesting that the protein is present in aggregates or inclusion bodies inside the bacterial cells. The molecular weights of the proteins did not correspond to the molecular weights at which the proteins' bands appeared in the gel, which can be explained by their isoelectric point, corresponding to 6.59, an acidic pI. Therefore, the proteins did not migrate as predicted, due to impairment of SDS binding [96]. Furthermore, the bands in the gel appeared at higher molecular weights than expected, possibly due to the proteins not undergoing complete denaturation, causing them to exhibit a non-linear conformation. However, at this stage, we did not pursue further purification due to time limitations.

3.2. S100B-Myc dimer and purification

To make S100B readily available for use in cell cultures and easily identifiable, we considered using a modified version of the S100B protein that could be detected using a specific marker. For the purpose we used a Myc tagged S100B construct, available at the host lab. The purification of S100B has been previously described, as well as the Myc tagged S100B form [90]. For the first and prime step of the purification of S100B-Myc, a hydrophobic interaction chromatography was performed. The usage of this technique takes advantage on the fact that S100B-Myc, just as other proteins of the S100 family, undergoes conformational changes upon Ca^{2+} binding through Ca^{2+}-binding sites in the EF-hand structural motifs [71]. Thus, upon binding to Ca^{2+}, the protein exposes hydrophobic regions to the solvent enabling the binding to the column matrix. For this reason, after cell lysis, the protein was incubated with $CaCl_2$ prior to injection to the PhenylSepharose6 column and eluted by switching the elution buffer to one containing EDTA without gradient. EDTA is a chelating agent that binds to divalent ions such as Ca^{2+}, leading to a conformational change on S100B-Myc and allowing its detachment from the matrix and further elution, as shown by the chromatogram and SDS-PAGE gel obtained in **Figure 3.2**.

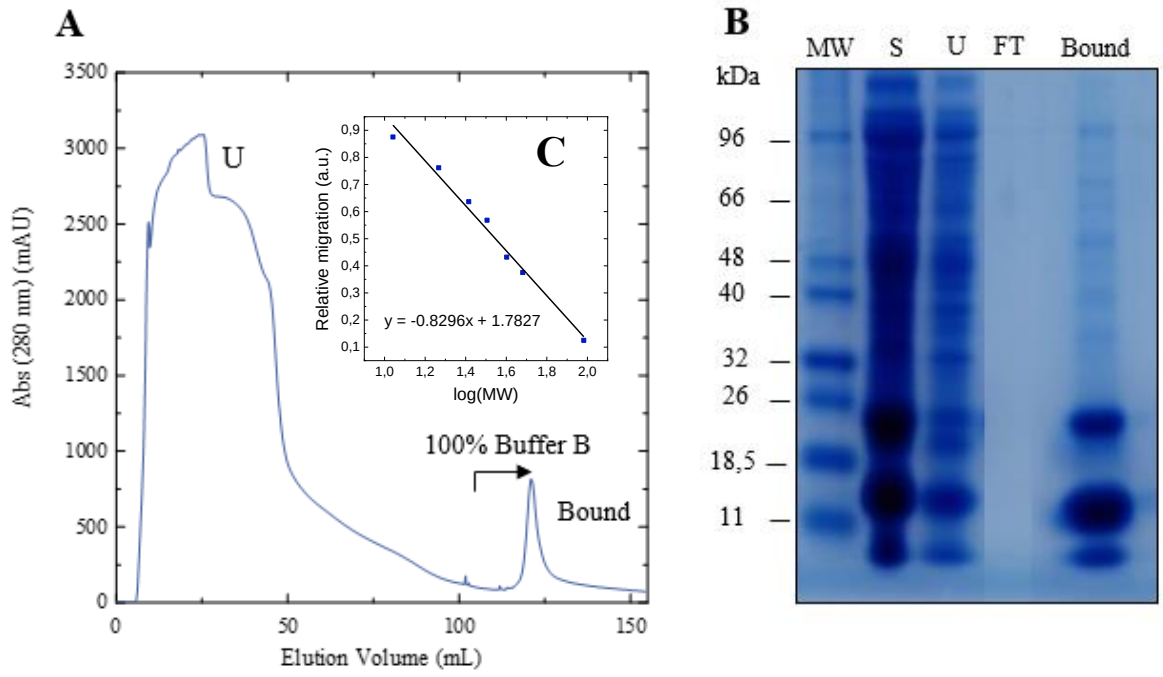

Figure 3.2 – S100B-Myc purification by hydrophobic interaction chromatography. (A) Hydrophobic interaction chromatogram (HiPrep Phenyl FF): U – unbound fraction; A – elution peak after elution buffer switch (S100B-Myc). **(B)** SDS-PAGE 8% Tris-Tricine of the chromatography elution peaks: MW – molecular weight standard (LMW II marker, NZYTech); S – supernatant after ultracentrifugation; U – unbound fraction; FT- flow-through fraction of the concentration with a 3 kDa cut-off membrane; Bound- peak fractions after concentration. **(C)** Calibration curve obtained for the SDS-PAGE.

The chromatogram obtained for the first step of S100B purification shows that the S100B-Myc fraction has a unique peak corresponding to the elution of the protein. SDS-PAGE analysis revealed the existence of 2 bands of overexpressed proteins. It was hypothesized that the high molecular weight band might correspond to S100B-Myc dimer and the lower molecular weight band to the protein monomer The analysis of the calibration curve obtained for the SDS-PAGE (**Figure 3.2C**) allowed to obtain a mass for each appearing band, corresponding to 21,9 kDa and 12,4 kDa, which comparing to the dimer mass (28,6 kDa) shows that the protein migrated more than expected, whereas comparing to the monomer (14,3 kDa), it appears as an approximated value. It is known that S100B migrates more than its apparent size, due to being an acidic protein [72] and also for the fact that it isn't completely denaturated as already stated, so it does not exhibit a linear profile. As the protein does not fully denature using the regular procedure of the SDS-PAGE technique, it is known that the protein can be quite stable and consequently, to some extent, resistant to denaturation.
The S100B-Myc fraction eluted from the PhenylSepharose6 column was concentrated and subjected to a size exclusion chromatography (SEC) in a Superdex 75 16/600 column ($V_{column} \approx 130$ mL) (**Figure 3.3**).

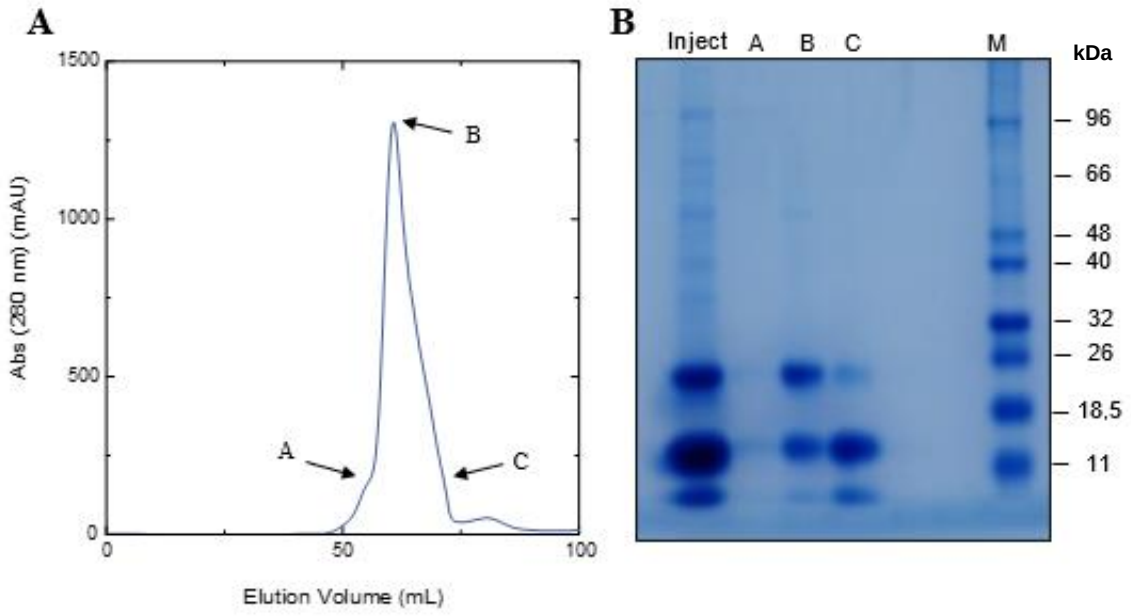

Figure 3.3 – S100B-Myc purification by size-exclusion chromatography. (A) Size-exclusion chromatogram (Superdex 75 16/600, HiLoad): A – first fraction of the elution peak; B – second fraction of the elution peak; C – third fraction of the elution peak. **(B)** SDS-PAGE 8% tris-tricine of the chromatography elution peaks: Inject – injected sample on SEC column; A – first fraction of the elution peak; B – second fraction of the elution peak; C – third fraction of the elution peak; M- Molecular weight marker (LMW II marker, NZYTech).

The separation by size exclusion chromatography revealed the existence of one peak corresponding to S100B-Myc elution. Analysis by SDS-PAGE of the dimer peak fractions showed once again not only the appearance of a band corresponding to the monomer fraction, as expected, but also another band with approximately the double of molecular weight of the monomer, probably the dimer that did not fully denaturate.

S100B binds with high affinity to Ca^{2+} and Zn^{2+} metal ions through the EF-hand sites or the interfacial site between dimers, respectively, by the presence of free cysteines that potentiate cross-linking bonding. This binding then induces conformational changes on S100B structure and can potentially increase the binding affinity of S100B (and S100B-Myc) to other proteins [97]. For this reason, an apo form of the protein (dimeric form without metal ions) was obtained through incubation with dithiothreitol (DTT)

and EDTA for 2 hours at 37 ºC followed by a size exclusion chromatography. DTT was used to reduce the disulfide bonds in S100B-Myc and promote the appearance of only dimeric species (not multimeric forms) and EDTA was used to act as a chelating agent to bind all the divalent ions present. After incubation, a size exclusion chromatography was performed using a Superdex 75 Tricorn column (Vc ≈ 26 mL) (**Figure 3.4**). The fraction containing S100B-Myc dimer was eluted as a symmetric band at 11.71 mL. The molecular weight was calculated using a calibration curve (**Figure 3.4C**), corresponding to 54.9 kDa, which does not correspond to the predicted molecular weight of the protein dimer (28.6 kDa). This difference might be correlated to the fact that S100B is not a globular protein, so the linear behaviour that underlies in the calibration curve cannot be applied to this protein, and the additional Myc tag might increase the hydrodynamic radius of the protein and then lead to an early elution, consequently to an overestimation of the molecular weight of S100B-Myc as inferred by the calibration curve.

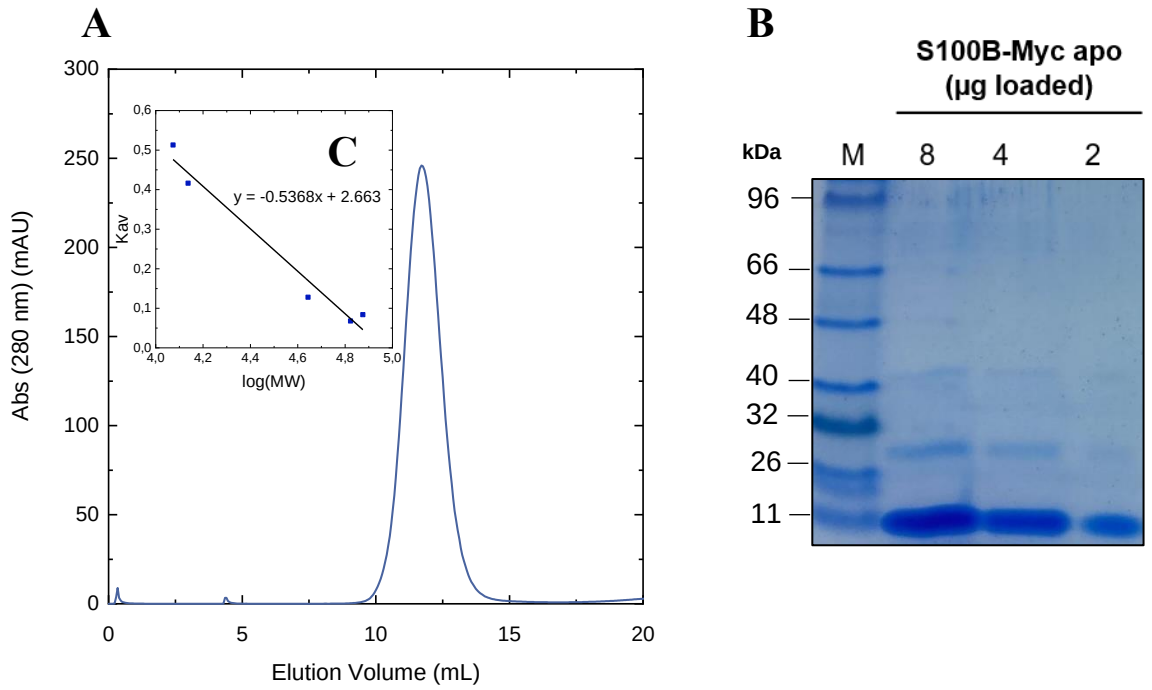

Figure 3.4 – Purification of S100B-Myc in the apo state by size-exclusion chromatography. (A) Size-exclusion chromatogram (Superdex 75). **(B)** SDS-PAGE 8% Tris-Tricine of the chromatogram elution peak: M – molecular weight marker (LMW II marker, NZYTech). The protein was loaded at different amounts (8, 4 and 2 µg). **(C)** Calibration curve obtained for the column used (Superdex 75 Tricorn).

The fraction containing S100B-Myc was eluted as a symmetric band shown in **Figure 3.4**. Purity was then confirmed by SDS-PAGE of the collected fraction. Interestingly, the apo state of the protein obtained from the size-exclusion chromatography exhibited a strong band near the predicted molecular weight calculated for S100B-Myc (14,3 kDa). This result confirmed the successful isolation of a highly pure form of S100B-Myc dimer.

3.3. Biophysical and biochemical characterization of S100B-Myc

3.3.1. Circular Dichroism spectroscopy

Circular dichroism (CD) is an essential analytical technique used to analyze chirality in molecules through their optical activity, thus its basis lies in the chiral nature of molecules, which describes the property of a molecule that cannot be overlayed onto its mirror image. The CD technique refers to the differential absorption of the left- and right-circularly polarized components of plane polarized electromagnetic radiation and occurs when a chromophore – an absorbing group – is chiral (optically active) for one or more of the following reasons: 1) it contains one or more chiral centers; 2) it is covalently bond to a chiral center; 3) it is placed in an asymmetric environment (asymmetric arrangement of substituents) (**Figure 3.5A**). [98]

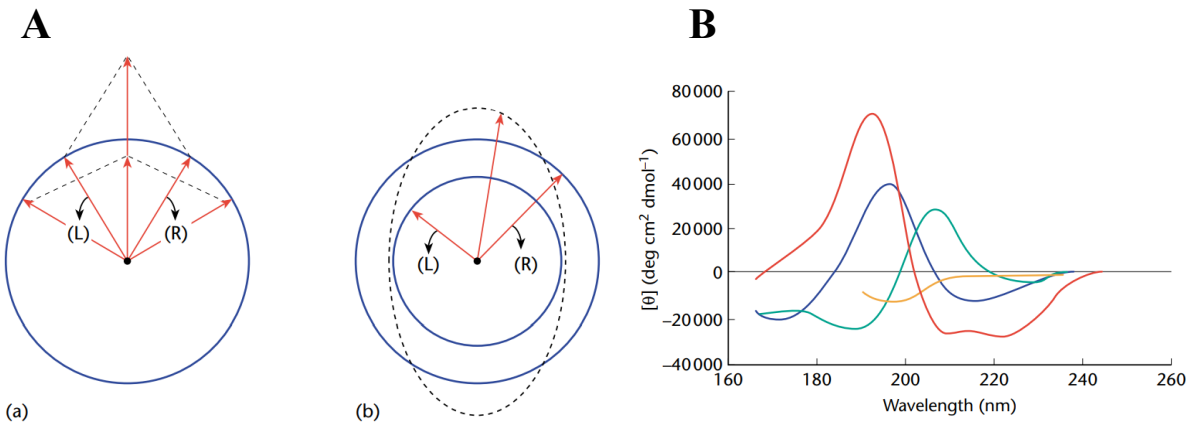

Figure 3.5 – Principles of circular dichroism spectroscopy. (A) Origin of the CD effect. The left- (L) and right- (R) circularly polarized components within plane-polarized radiation: **(a)** When these two components have the same amplitude and when combined, they yield plane-polarized radiation; **(b)** If these components are of different magnitudes, the resultant light (depicted by the dashed line) is elliptically polarized. **(B)** The far-ultraviolet (UV) circular dichroism (CD) spectra associated to diverse secondary structure elements found in proteins. The CD spectra are color-coded for clarity: red represents α-helix, blue represents antiparallel β-sheet, green corresponds to type I β-turn, and orange denotes irregular or random coil structure. [99]

The CD spectrum of the far-UV region (typically ranging from 240 nm to 190/180 nm can be used to provide quantitative estimates of the secondary structure of a protein. In this region of the spectrum the absorbing chromophore is essentially the peptide bond, with a weak though broad transition ($n \rightarrow \pi^*$) around 210 nm and an intense transition ($\pi \rightarrow \pi^*$) at about 190 nm. The different forms of regular secondary structure found in peptides and proteins (α-helices, β-sheets, random coils) exhibit distinct far-UV CD spectra, therefore is possible to infer the secondary structure composition of a protein by comparing experimental CD spectra with reference spectra. It is possible to make estimates of the α-helical content from the ellipticities at 208 and 222 nm, that correspond to the double minima characteristic of the α-helix signature (**Figure 3.5B**). [98]

To study the biophysical characteristics of the S100B-Myc protein, far-UV circular dichroism spectroscopy assays were employed to investigate the secondary structure elements of the S100B protein. As a first approach, the far-UV spectra of the protein were obtained, both in the holo (calcium-bound state) and in the apo forms of the protein. Previous studies have already shown that both S100B

and the S100B-Myc tagged have a predominance of α-helical secondary structure [90], but the last with no distinction between the holo and apo forms, which will be addressed in this study. In **Figure 3.6A**, the far-UV spectra of the holo form of the protein at 20ºC, 97ºC and 20ºC recovery spectra are shown. At 20ºC, the protein exhibits a secondary structure predominance of α-helix signature, with 2 characteristic minima at 208 and 222 nm typical of an α-helical conformation. When the temperature increases to 97ºC, the protein still shows two minima at the same wavelengths as before, and the α-helical structure is still prominent, which might suggest that the protein does not fully denaturate at this temperature. By observing the recovery spectra taken at 20ºC after the thermal ramp, it is possible to infer that the protein partially refolds into the initial secondary structure.

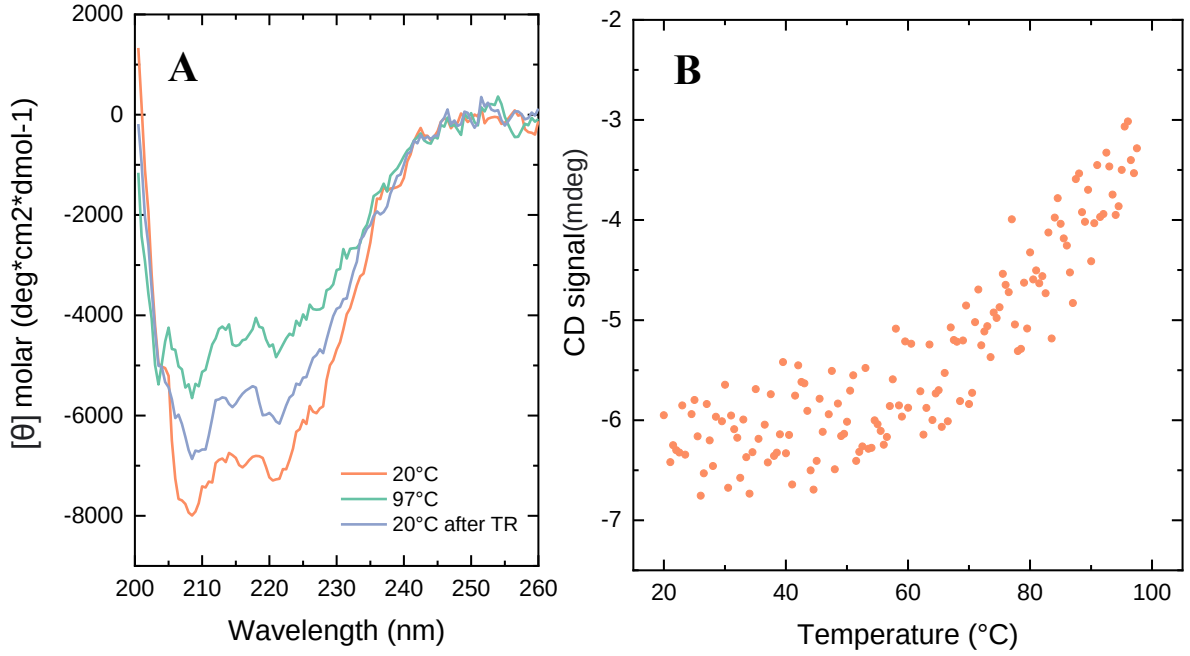

Figure 3.6 – Far-UV circular dichroism measurements of S100B-Myc holo. (A) Temperature-dependent far-UV CD spectra of S100B-Myc protein in the holo state: A comparative depiction of CD spectra acquired at 20ºC, 97ºC and 20ºC after a thermal ramp from 20°C to 97°C (recovery spectra). The initial and final 20°C spectra serve as references of the protein in its native holo state, while the 97°C spectra and the 20ºC recovery spectra offers insight into the unfolding and refolding behavior upon thermal stress, respectively. S100B-Myc does not fully denaturate at 97ºC, and the recovery spectra shows that the protein partially recovers its secondary structure. **(B)** Thermal denaturation profile of S100B-Myc protein: Temperature was gradually increased from 20ºC to 97ºC, at a heating rate of 1ºC/min, revealing the protein response to thermal stress. The absence of a distinct melting temperature (Tm) suggests exceptional stability under the tested conditions.

To investigate the thermal denaturation behavior of the protein, a thermal ramp experiment was conducted (**Figure 3.6B**). In this study, the protein sample was subjected to a controlled increase in temperature, from 20ºC to 97ºC, at a heating rate of 1ºC/min. This technique allowed the systematic observation of how the protein structural integrity changes as a function of temperature. In other words, as the temperature increases, changes in the protein secondary, tertiary, and quaternary structure were monitored. By monitoring the changes of the protein with the increase of temperature, it's possible to infer about protein's stability as well as its unfolding patterns. The thermal denaturing curve obtained from a thermal ramp experiment generally follows a sigmoidal-like shape, from which it is possible to extract the melting temperature (Tm), which corresponds to the temperature at which the protein undergoes a significant structural change (critical point in which 50% of the protein population undergoes unfolding), such as denaturation. The Tm is typically estimated as the midpoint of the sigmoidal transition. However, from the results presented, the absence of a distinct Tm in the thermal ramp of S100B-Myc in the holo state strongly suggests that the protein remained stable under the thermal

denaturing conditions. The fact that the protein fails to exhibit a clear Tm, suggests that the unfolding process did not occur within the investigated temperature range. This outcome underscores the protein's high stability, indicating its ability to withstand elevated temperatures without undergoing substantial structural alterations. This result corroborates what was mentioned before, when the protein was loaded into an SDS-PAGE gel (also under denaturing conditions) did not monomerize as expected, enhancing the protein's stability and resistance to the destabilizing conditions applied. The same protocol was followed for the S100B-Myc in the apo state. The CD spectra and the thermal ramp obtained are shown in **Figure 3.7**.

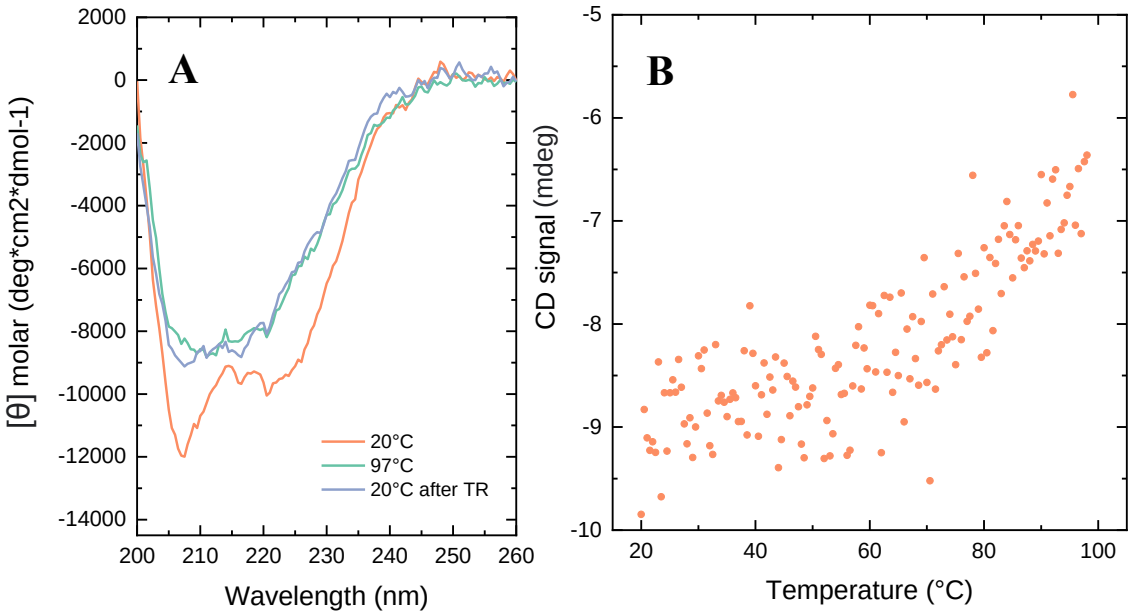

Figure 3.7 – Far-UV circular dichroism measurements of S100B-Myc apo. (A) Temperature-dependent far-UV CD spectra of S100B-Myc protein in the apo state: A comparative depiction of CD spectra acquired at 20ºC, 97ºC and 20ºC after a thermal ramp from 20°C to 97°C (recovery spectra). The initial 20°C spectra serves as reference of the protein in its native apo state, while the 97°C spectra and the 20ºC recovery spectra offers insight into the unfolding and refolding behavior upon thermal stress, respectively. **(B)** Thermal denaturation profile of S100B-Myc protein: The temperature was gradually increased from 20ºC to 97ºC, revealing the protein's response to thermal stress. The absence of a distinct melting temperature (Tm) suggests exceptional stability under the tested conditions.

For the apo state of the S100B-Myc protein, the recovery spectra recorded at 20ºC after the thermal denaturation appears very similar to the spectra captured at 97ºC. This similarity indicates that there is no significant refolding of the protein after exposure to the elevated temperatures. Furthermore, the thermal ramp experiment, which involved a gradual temperature increase from 20ºC to 97ºC, has once again failed to manifest a distinctive melting temperature in the resulting melting curve. This recurrence of the absence of a discernible Tm in the thermal ramp reaffirms the protein's inherent stability under the tested conditions, corroborating the results obtained of the holo state. The combined results obtained from these spectral and thermal analyses underscores the high stability of the S100B-Myc protein, emphasizing its capacity to maintain its structural conformation even when subjected to high temperature environments.

3.3.2. ANS fluorescence emission

To proceed with the biophysical characterization of S100B-Myc, ANS assays were performed to uncover the intricate aspects of the protein's structural behavior and dynamics upon ligand binding.

ANS – 8-anilino-1-naphthalenesulfonic acid – is an extrinsic fluorescent probe that exhibits distinct fluorescence patterns upon binding to hydrophobic regions within proteins. When bound to buried hydrophobic sites of proteins, ANS exhibits a characteristic blue shift (or hypsochromic shift) of fluorescence emission maxima and an increase of fluorescence intensity. ANS binds noncovalently to proteins and its fluorescence varies with changes in the solvent environment.

This fluorescent dye enables the study of conformational changes occurring in proteins upon the addition of ligands that induce these structural alterations, as well as also be used for the evaluation of folding/unfolding processes within proteins. [100] [101]

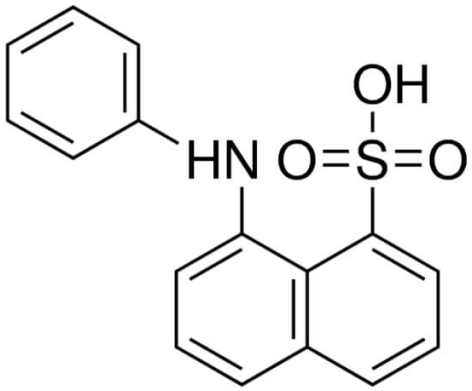

Figure 3.8 – Chemical Structure of ANS (8-Anilino-1-Naphthalenesulfonic Acid). The image illustrates the molecular composition and arrangement of ANS, a fluorescent probe known for its distinct fluorescence behavior upon interaction with hydrophobic regions of proteins.

Therefore, ANS measurements are a valuable tool for the detection and analysis of conformational changes in proteins. As previously noted, S100B is a Ca^{2+}-binding protein, which binding via EF-hand motifs induces conformational changes that underlie some of the functional interactions with other proteins [90]. To study these structural modifications on the S100B-Myc tagged protein, an ANS spectroscopy assay was performed using the apo form of the protein.

The ANS fluorescence emission spectra of S100B-Myc were analyzed both in the absence and presence of Ca^{2+} ions. Two distinct Ca^{2+} concentrations were employed: 1 mM and 2 mM (excess). Notably, the addition of both 1 mM and 2 mM Ca^{2+} caused a progressive and marked blueshift in the fluorescence emission peak, indicating significant structural alterations within the protein upon Ca^{2+} addition. Moreover, an evident enhancement in fluorescence intensity was also observed upon Ca^{2+} addition, suggesting changes in the local environment surrounding hydrophobic regions with the increment of Ca^{2+} concentration **Figure 3.9**.

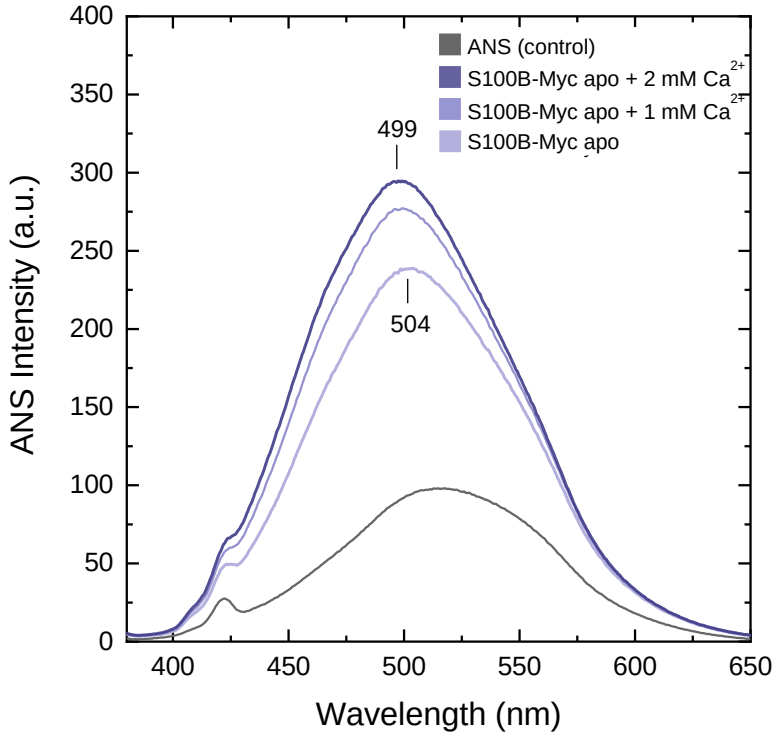

Figure 3.9– ANS fluorescence emission spectra in the absence and presence of Ca^{2+}. Ca^{2+} was added in two different concentrations, 1mM and 2mM (excess) and S100B-Myc apo was in a final concentration of 10 µM. Both the addition of 1 mM and 2 mM of Ca^{2+} exhibit a pronounced blueshift in the fluorescence emission, as well as a progressive increase in fluorescence intensity.

3.4. Analysis of Htt aggregation in the presence of S100B by the FRASE assay

The FRASE assay – FRET-based mHtt aggregate seeding (FRASE) assay – is a technique that enables the quantification of mHtt seeding activity in complex biosamples extracted from HD patients and disease models. In the context of this study, the FRASE assay was employed with the aim of studying the effect of S100B in huntingtin aggregation.

In the FRASE assay, two soluble purified glutathione S-transferase (GST)-tagged HttExon1 proteins with a polyQ tract length of 48 glutamines were C-terminally fused with two reporter proteins, CyPet and YPet – (GST-HttEx1Q48-CyPet and GST-HttEx1Q48-YPet). The two constructs were only soluble in aqueous solutions as long as they contained the GST tag in the N-terminal. Upon cleavage of the GST tag using the PreScission protease (PP), the HttEx1Q48-CyPet and -YPet fragments were released and initiated spontaneous co-aggregation. Their co-aggregation could be quantified by measuring FRET as the fluorescent tags CyPet and YPet were brought into close proximity when fibrillar mHttEx1 aggregates were formed, as represented in **Figure 3.10**. [91]

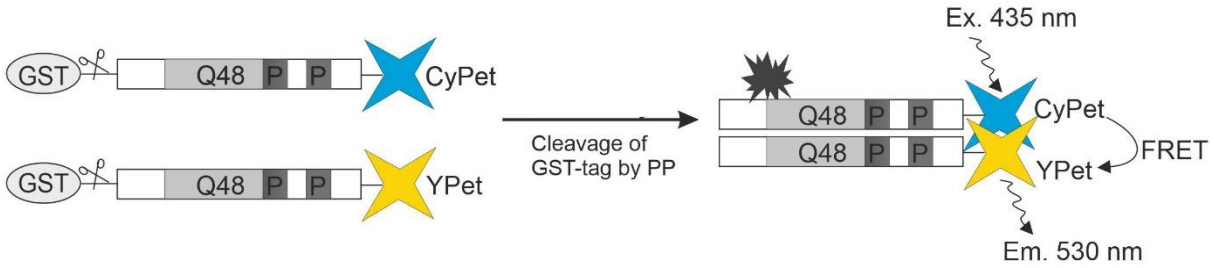

Figure 3.10 – Schematic representation of FRET-based mHtt aggregate seeding assay (FRASE) for testing protein-protein interactions. Co-aggregation of HttEx1Q48-CyPet and HttEx1Q48-YPet is induced upon cleavage of GST fusion proteins with PreScision Protease (PP) and it could be quantified by measuring FRET as the fluorescent tags CyPet and YPet are brought into close proximity when aggregates are formed. (Adapted from [91])

To investigate the influence of S100B in huntingtin aggregation, the first approach was to test different concentrations of S100B, ranging from 0.5 µM, 1 µM, 2 µM, to 8 µM, while maintaining a fixed concentration of 2 µM for HttExon1-Q48 (CyPet and YPet fused). In all experimental sets, there was no addition of Ca^{2+} ions.

QBP1 is a peptide that inhibits the aggregation of polyglutamine (polyQ) proteins [102], so it was used as a positive control. In **Figure 3.11**, the plot shows that across all S100B-added conditions and their respective concentrations, there was no difference on huntingtin aggregation observed, whereas in the situation where QBP1 was added, there was a evident delay on the process of aggregation, as expected. These results suggest and might corroborate the fact that S100B chaperone activity is induced by the presence of Ca^{2+} ions.

.

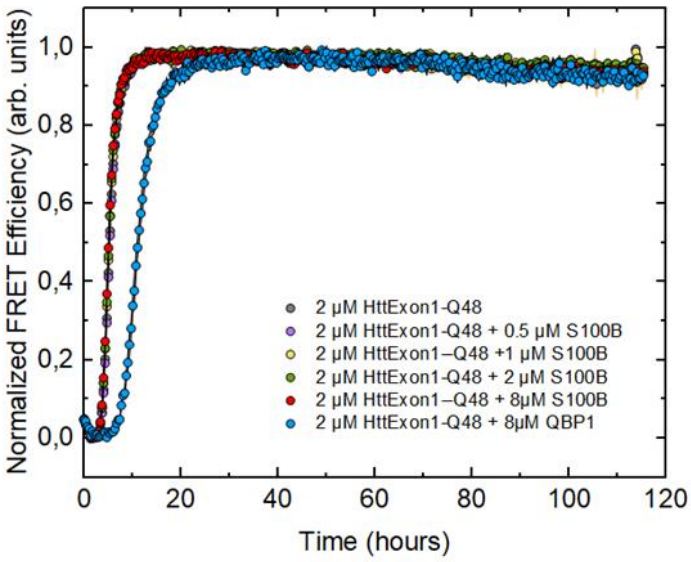

Figure 3.11 – Investigating S100B's impact on huntingtin aggregation by the FRASE assay. The plot presents the assessment of the influence of different concentrations of S100B on huntingtin aggregation dynamics. Concentrations of S100B spanning from 0.5 µM, 1 µM, 2 µM, to 8 µM were tested, while maintaining a fixed concentration of 2 µM for HttExon1-Q48 (CyPet and YPet fused). As a comparative benchmark, the aggregation inhibitor peptide QBP1 was utilized as a positive control (blue scatter). The plot demonstrates that, across all S100B-added conditions and concentration ranges, there was no apparent impact on the aggregation of huntingtin. In contrast, the addition of QBP1 exhibited a delay in the aggregation process.

Since S100B's chaperone-like function was described to be calcium dependent [79, 80], a second FRASE assay was subsequently conducted, testing the effect of calcium bound-S100B on huntingtin aggregation. All the results are shown in **Figure 3.12**.

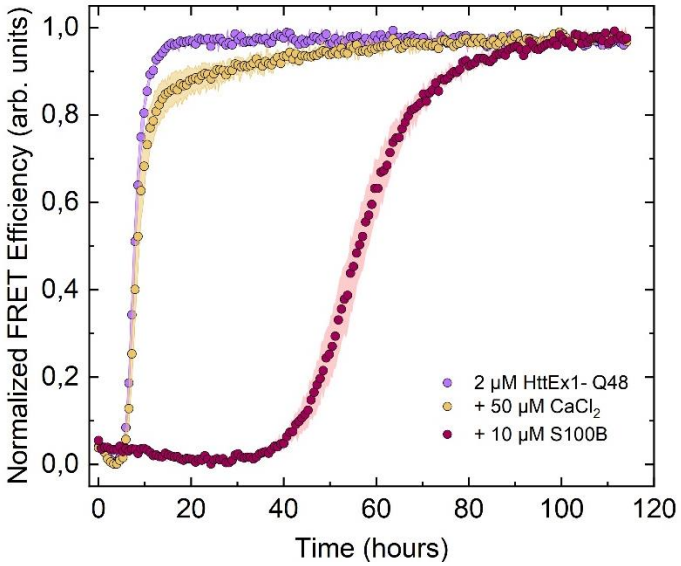

Figure 3.12 – S100B-Ca^{2+} inhibits huntingtin aggregation monitored by the FRASE assay. Kinetic traces of FRET-monitored aggregation of HttEx1Q48-CyPet and -YPet fragments after GST cleavage. In purple, 2 µM of both constructs HttEx1Q48-CyPet and -YPet display a very rapid aggregation. In yellow, 2 µM of both constructs HttEx1Q48-CyPet and -YPet in the presence of 50 µM of Ca^{2+} show that the aggregation profile of huntingtin displayed remarkable similarity to the initial state, suggesting a limited influence of Ca^{2+} on the aggregation kinetics. In red, 2 µM of both constructs HttEx1Q48-CyPet and -YPet in the presence of 50 µM of Ca^{2+} and with the incorporation of 10 µM of S100B, shows a significant and evident delay in the aggregation process of huntingtin in the presence of calcium.

As **Figure 3.12** shows, huntingtin aggregation occurred within approximately 20 hours. However, when S100B-Ca^{2+} was present, in suprastoichiometric concentrations, an evident delay in the huntingtin aggregation process was observed, extending the aggregation time to about 100 hours, as indicated by the red scatter data in **Figure 3.12**. To ensure that this effect was specific to S100B and not due to calcium alone, a control condition was tested, by adding calcium to huntingtin. This result revealed that calcium alone does not impact huntingtin aggregation. These findings emphasize the significant role of S100B in modulating huntingtin aggregation dynamics. Furthermore, our study suggests that S100B may serve as a novel regulatory factor in controlling protein aggregation within the context of Huntington's disease (HD), offering promising avenues for further investigation.

3.5. Investigation Htt and S100 proteins interactions resorting to the LuTHy assay

To test direct interactions between huntingtin and S100B, a LuTHy assay was performed. LuTHy is a bioluminescence-based two-hybrid method that enables the detection of binary protein-protein interactions (PPIs) with high sensitivity and specificity in mammalian cells, through the combination of two readouts in one experiment. [92]

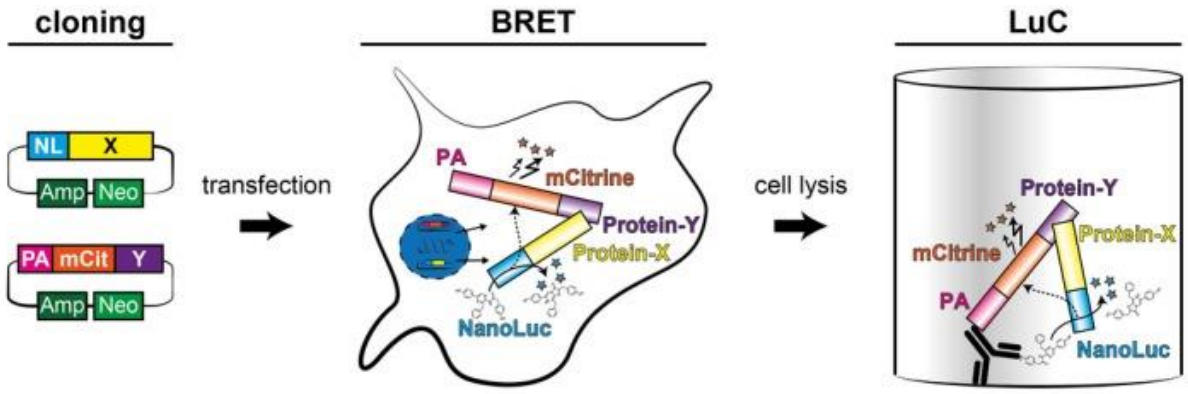

Figure 3.13 – Schematic representation of the procedure of the LuTHy assay. Expression vectors encoding NL and PA-mCit-tagged hybrid proteins are cloned and co-transfected into mammalian cells. Binary interactions are detected with a double readout; first, in vivo with in-cell bioluminescence resonance energy transfer (BRET) and second, ex vivo with a luminescence-based co-precipitation (LuC). Protein-X and Protein-Y represent the two proteins of interest whose interactions are being studied. (Adapted from [92])

This experiment required the construction of two distinct hybrid fusion proteins, one tagged to NanoLuc luciferase (NL) and another to ProteinA-mCitrine (PA-mCit). These tags could be positioned N-terminally or C-terminally in the hybrid fusion proteins, that comprised huntingtin and one of its three studied interactors: FEZ1, S100A1 and S100B. In the context of this work, S100A1 was used as a negative control. Conversely, prior studies have reported that FEZ1 binds to huntingtin, so within this experiment it was used as a positive control [103] [104]. The interactions between the pairs of proteins of interest were detectable in cells (*in vivo*) through quantification of bioluminescence resonance energy transfer (BRET). Bioluminescence resonance energy transfer (BRET) is a transfer of energy between a luminescence donor – NanoLuc (NL), upon addition of its substrate (in this case, Coelenterazine) and a fluorescence acceptor – mCitrine, upon excitation of the energy transferred by the donor. The technology involves the fusion of donor and acceptor molecules to proteins of interest. Co-expression of the fusion constructs in living cells enables their interaction to be studied in real time in a quantitative manner [105]. This occurs when the luciferase-tagged donors and mCitrine-tagged fluorescent acceptors come into close proximity, at a distance inferior to 10 nm. Protein complex formation between NL-X and PA-mCit-Y (being X and Y the two proteins of interest) should subsequently be detectable in lysed cells (*ex vivo*) with a bioluminescence-based co-precipitation (Luc) assay. In this assay, the PA-tag in the hybrid protein is used for bait precipitation, while the NL-tag is used for the detection of the interacting prey. The LuTHy principle is shown in **Figure 3.13**. [92]

To assess the interactions among the proteins of interest (Htt and FEZ1, Htt and S100A1, and Htt and S100B), all potential orientations of the tagged NL and PA-mCitrine were tested. The variety of experimental configurations aimed to investigate the full spectrum of possible interactions. **Table 3.1** provides an overview of all the potential interactions investigated in this experimental framework.

Table 3.1 – Overview of the protein-protein interactions studied with distinct orientations in the LuTHy assay. The NL-tagged proteins act as a donor and PA-mCit tagged proteins act as acceptors for BRET measurements. When the interactors act like donors (NL-tagged), huntingtin act as an acceptor (PA-mCit tagged), and vice-versa. The interactions were tested in all possible orientations with three interactors: FEZ1 (positive control), S100A1 (negative control) and S100B (protein of interest).

	FEZ 1	S100A1	S100B	
Interactor as a donor	FEZ1-NL + Htt-mCit-PA	S100A1-NL + Htt-mCit-PA	S100B-NL + Htt-mCit-PA	N-terminally tagged Huntingtin
	NL-FEZ1 + Htt-mCit-PA	NL-S100A1 + Htt-mCit-PA	NL-S100B + Htt-mCit-PA	
	FEZ1-NL + PA-mCit-Htt	S100A1-NL + PA-mCit-Htt	S100B-NL + PA-mCit-Htt	C-terminally tagged Huntingtin
	NL-FEZ1 + PA-mCit- Htt	NL-S100A1 + PA-mCit-Htt	NL-S100B + PA-mCit-Htt	
Interactor as an acceptor	FEZ1-PA-mCit + NL-Htt	S100A1-PA-mCit + NL-Htt	S100B-PA-mCit + NL-Htt	N-terminally tagged Huntingtin
	PA-mCit-FEZ1 + NL-Htt	PA-mCit-S100A1 + NL-Htt	PA-mCit-S100B + NL-Htt	
	FEZ1-PA-mCit + Htt-NL	S100A1-PA-mCit + Htt-NL	S100B-PA-mCit + Htt-NL	C-terminaly tagged Huntingtin
	PA-mCit-FEZ1 + Htt-NL	PA-mCit-S100A1 + Htt-NL	PA-mCit-S100B + Htt-NL	

For a comprehensive approach, the investigation studied the interactions between the three targeted interactors (FEZ1, S100A1 and S100B) and seven distinct constructs of huntingtin, that are shown in **Table 3.2**. The diversity of the constructs used allowed to broaden the scope of exploration into the interactions between huntingtin and the studied interactors, facilitating a thorough assessment of the PPIs studied.

Table 3.2 – Huntingtin constructs investigated in protein-protein interaction analysis by the LuTHy assay. Seven distinct huntingtin constructs were analyzed to investigate their interactions with the target interactors (FEZ1, S100A1, and S100B).

Huntingtin constructs	HttExon1-Q23
	HttExon1-Q79
	Htt(1-513)-Q23
	Htt(1-513)-Q43
	Htt(1-513)-Q79
	Htt full length-Q23
	Htt full length-Q145

The output of the experiment consists of both BRET and Luc values measured for each potential interaction. The BRET values detected were further corrected (cBRET values), and according to the literature, "true" interactions were considered with a cBRET value above or equal to 0.01. The cLuc (corrected Luc) values obtained in this experiment were all approximately 0, therefore considered non-contributory to the analysis, so the discussion of these data was discarded. In **Figure 3.14**, the cBRET values measured and obtained for all the possible orientations between the HttExon1-Q23 construct and the interactors tested are represented in a violin plot.

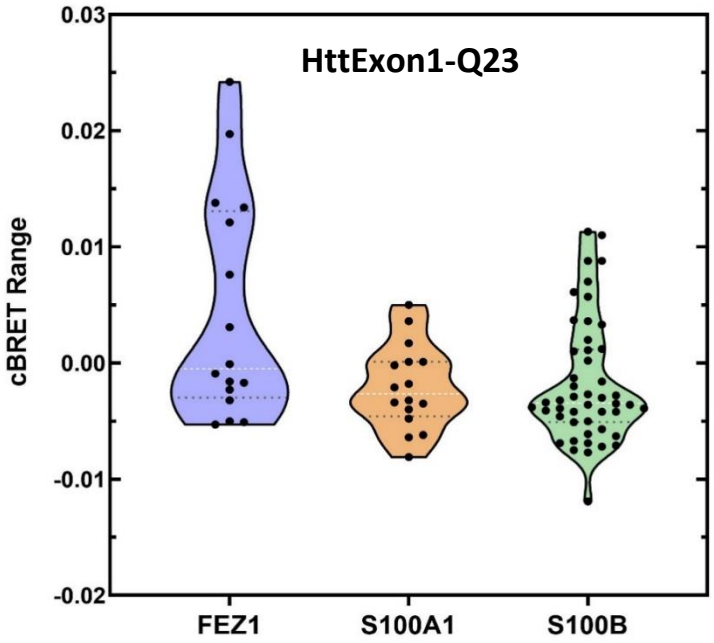

Figure 3.14 – Violin plot of cBRET values for interactions between HttExon1-Q23 and FEZ1, S100A1, and S100B. The violin plot illustrates the cBRET values obtained from the interactions involving HttExon1-Q23 and the tested possible interactors: FEZ1, S100A1, and S100B. The cBRET values, serving as indicators of potential protein-protein interactions, are represented along the y-axis, while the x-axis depicts the distinct potential interactors. Upon examination of the plot, it is evident that the majority of cBRET values fall below the established threshold for confirming positive interactions (cBRET≥0.01).

From the analysis of **Figure 3.14**, it is inferable that most of the data points presented are below the threshold set for confirming a positive protein-protein interaction. A noteworthy observation is the distinct behavior of the positive control, FEZ1, which is marked by an absence of values surpassing the predefined cutoff (≥0.01). Given that the positive control did not yield the expected outcomes, it raises the potential concern about the reliability of all other measurements obtained in the experiment.

The remaining violin plots for all other constructs are collectively presented in **Figure 3.15**. From **Figure 3.15** the overall data tends to fall below the established threshold. However, some constructs exhibited more relevant results. For the HttExon1-Q79, the interactions study with the positive control FEZ1 have some orientations above the threshold, as well as some of the interactions with S100B, which might suggest that this particular construct establishes some relevant interactions with the chaperone. Another construct that stands out is the Htt(1-513)-Q79, in which the positive control also has positive results (above the threshold) as well as the interactions studied with S100B. While not definitive, these results suggest that there might be an interaction between huntingtin and S100B, even though the assay requires optimization for enhanced accuracy and reliability of the results.

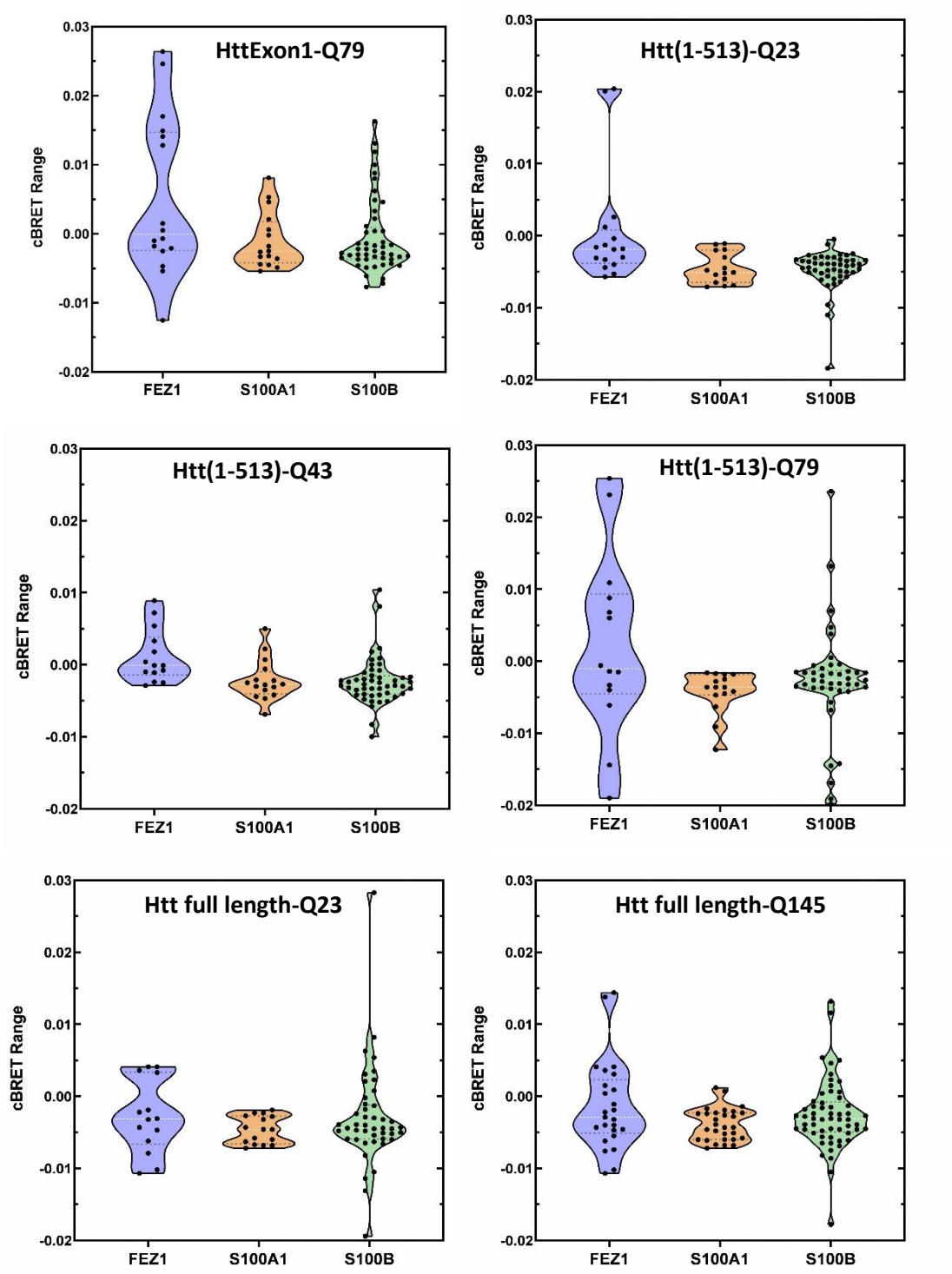

Figure 3.15 – Violin plots of cBRET values for interactions between Huntingtin and FEZ1, S100A1, and S100B for HttExon1-Q79, Htt(1-513)-Q23, Htt(1-513)-Q43, Htt(1-513)-Q79, Htt full length-Q23 and Htt full length-Q145. Some outliers were removed from the data set represented. It is evident that the majority of cBRET values present in the violin plots shown for all the HttExon1 constructs lie below the predefined threshold required for confirming positive interactions (cBRET≥0.01).

3.6. BiFC assay to study Htt aggregation in HeLa cells

The BiFC assay, or Bimolecular Fluorescence Complementation assay, is a technique used in molecular and cell biology to study protein-protein interactions within living cells. It allows the visualization and detection of the interaction between two target proteins by using fluorescence microscopy. The basic principle of the BiFC assay involves splitting a fluorescent protein into two non-fluorescent fragments, which are then individually fused to the proteins of interest. These protein fusions are co-expressed in the same cell, and if the two target proteins interact with each other, the two halves of the fluorescent protein are brought together, reconstituting and forming a functional fluorescent protein. This reconstituted fluorescent protein emits fluorescence when excited with appropriate wavelengths of light, which can be visualized and quantified using fluorescence microscopy. The fluorescent protein used in this assay was the Venus protein, an enhanced yellow fluorescent protein published in 2002, which is an *Aequorea victoria* GFP variant protein. The N-terminal of Venus consists of amino acid residues 1–158, referred to as Venus-N (VN), while its C-terminal encompasses residues 159–239, designated as Venus-C (VC). The principle of the BiFC assay performed is illustrated in **Figure 3.16.** BiFC assay provides a direct visualization of protein-protein interactions within the cellular context, allowing to study these interactions in real-time and in *vivo*, with minimal perturbation of the normal cellular environment. [106] [107]

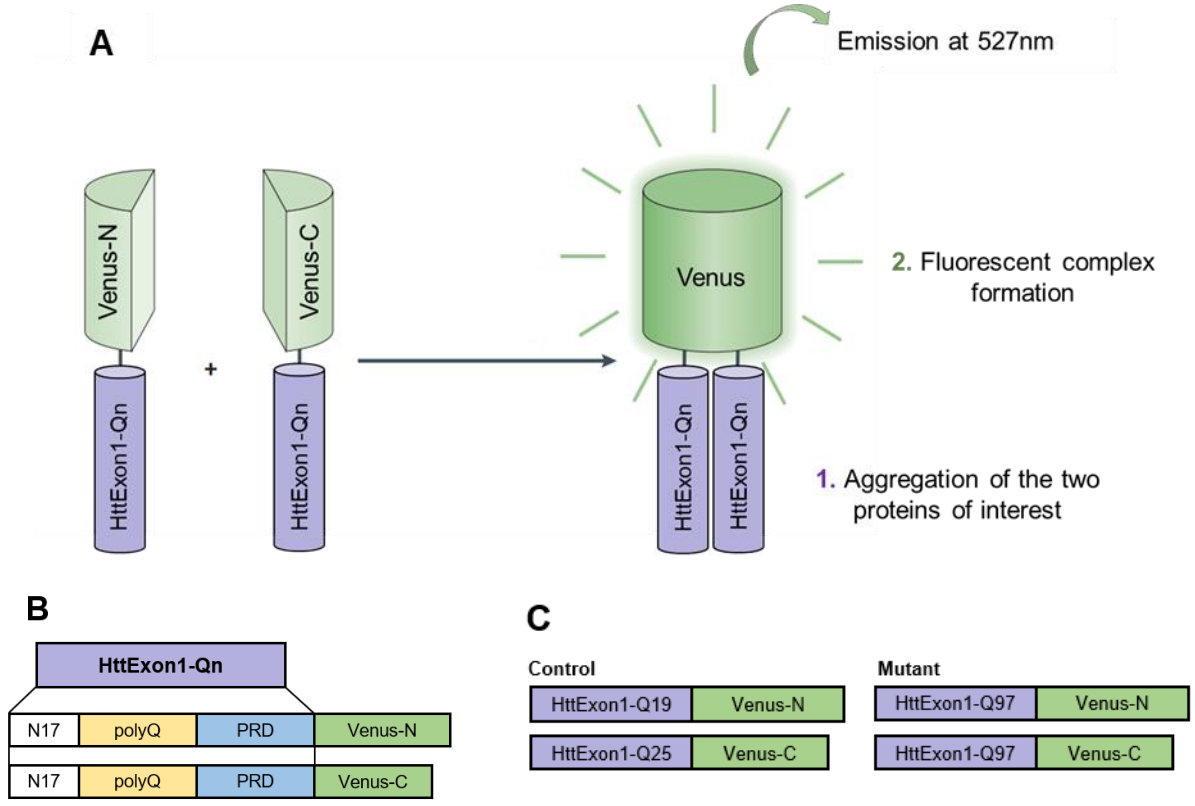

Figure 3.16 – Principle of bimolecular fluorescence complementation (BiFC) for HttExon1-Qn aggregation study. (A) The BiFC approach is based on the formation of a bimolecular fluorescent complex when two non-fluorescent fragments of the Venus protein are brought together by an interaction between HttExon1-Qn proteins (Adapted from [107]). **(B)** Schematic representation of the HttExon1-Qn BiFC constructs (Adapted from [108]). **(C)** Schematic outline of the different HttExon1-Qn constructs used. The first pair studied was considered the control pair (with a normal range of polyQ length) – pair HttExon1-Q19-N-Venus/HttExon1-Q25-C-Venus. The second pair HttExon1-Q97-N-Venus/HttExon1-Q97-C-Venus is referred to as the mutant pair, with an expanded polyQ tract.

3.6.1. Investigation of differential expression of Htt-Exon1-Q25/Q19 and Htt-Exon1-97Q/97Q in cells

To study the variations in expression levels and distinctive aggregation patterns between wild-type (control) and mutant HttExon1 fragments, bimolecular fluorescence complementation assays were employed. This approach enabled the investigation of the co-localization of HttExon1 fragments within the cellular space through live imaging. For this, the HeLa cell line was transfected with the two BiFC pairs: wild-type (HttExon1-Q19/HttExon1-Q25) and mutant (HttExon1-Q97/HttExon1-Q97). The transfection was executed with independent introduction of each pair into the HeLa cells, creating two distinct experimental contexts, one with HeLa cells expressing wild-type protein, and the second with HeLa cells expressing its mutant form. The differences were observed using fluorescence live-cell imaging microscopy, and the aggregation morphology of each BiFC pair was analysed. For the control pair, 0.5 µg of each plasmid was transfected, whereas for the mutant BiFC pair 1 µg of each plasmid was transfected. The difference of the amount of plasmids transfection between the wild-type and the mutant plasmids is due to the fact that with longer polyQ stretches, the cellular machinery undergoes atrophies and protein expression is not as efficient as with the expression of a normal range polyQ length. In an attempt to have comparable amounts of protein in the two experiments, the quantity of plasmid transfected for the mutant was double that of the control. In **Figure 3.17A** and **Figure 3.18**, the results of wild-type and mutant HttExon1 fragments are shown, respectively.

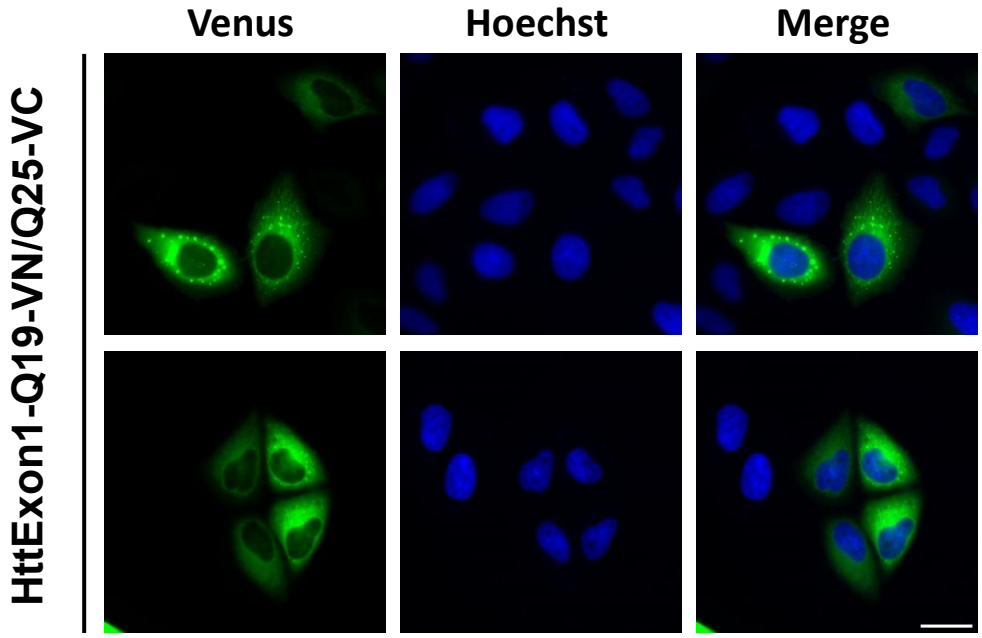

Figure 3.17 – Live-cell imaging of HeLa cells transfected with HttExon1-Q19/Q25 fragments in live HeLa cells using fluorescence microscopy. Cells were seeded on ibiTreat dishes and co-transfected with BiFC pair constructs for 48h. The Venus channel captures the fluorescence emitted by the Venus-tagged HttExon1-Q19/Q25 fragments, revealing their distribution and interactions within the cell. The Hoechst channel, represented in blue, highlights the nuclei of the cells providing insights on the distribution of the cell sand offering a point of reference for their localization (Hoechst 33342). The aggregation pattern exhibited by the HttExon1-19Q/25Q constructs is predominantly punctae, indicating the formation of discrete aggregation points and the fluorescence appears to be more broadly dispersed throughout the cell. Scale bar = 30 µm.

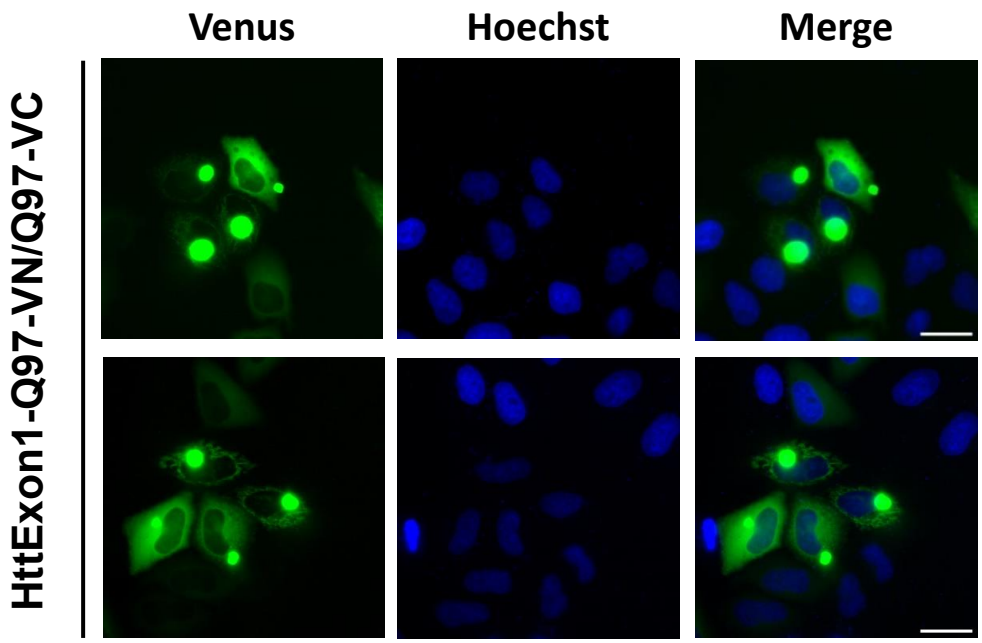

Figure 3.18 – Live-cell imaging of HeLa cells transfected with HttExon1-Q97/Q97 fragments in live HeLa cells using fluorescence microscopy. Cells were seeded on ibiTreat dishes and co-transfected with BiFC pair constructs for 48h. The Venus channel captures the fluorescence emitted by the Venus-tagged HttExon1-Q97/Q97 fragments, revealing their distribution and interactions within the cell. Nuclei were stained with Hoechst 33342. The aggregation pattern exhibited by the HttExon1-Q97/Q97 constructs appears mostly in aggresomes, large inclusion bodies present in the cytosolic region of the cell. Scale bar = 30 µm.

The wild-type transfected cells exhibited a dispersed fluorescence, notable for its even distribution through the cytoplasmatic region, with the presence of punctae aggregates, as observed by the representative examples showcased in **Figure 3.17**. These cells displayed the presence of large cytoplasmatic inclusions, with distinct structural arrangement compared to the wild-type transfected cells. The mutant form of the protein tends to accumulate and aggregate in inclusion bodies, though it may also appear with a punctae morphology, while the aggregates HttExon1 fragment containing a normal polyQ length did not generate large insoluble aggregates, which was already previously reported in the literature [109]. The distinct aggregation patterns between the wild-type and mutant constructs provide valuable understanding into the mechanisms shaping the behaviour of HttExon1 fragments with differing polyQ lengths.

The expression of the BiFC pairs was monitored by Western blot with the analysis of total cells lysates obtained after the protein extraction protocol. GAPDH (Glyceraldehyde-3-phosphate dehydrogenase) is one of the key enzymes involved in glycolysis, considered as a housekeeping protein. It is constitutively expressed in almost all tissues in high amounts, and for this reason, this protein was used as a loading control in the western blotting for the soluble fraction. For the insoluble fraction, the control used was actin, also a considered housekeeping protein involved in cell motility, structure, and integrity. [110]The corresponding bands for the wild-type and mutant HttExon1 fragments are shown in **Figure 3.19**, confirming the integration of both wild-type and mutant BiFC pairs into the cells after the transfection and its subsequent expression. The BiFC pairs were detectable both in the soluble and insoluble fractions, maybe due to Htt's even distribution in some cells, leading to its appearance in the soluble fraction, or due to the presence of the aggregates, whether punctae or aggresomes, that appear in the insoluble fraction. Also, Htt may interact with diverse cell compartments, which can alter its solubility and therefore explain the appearance on the two fractions analysed.

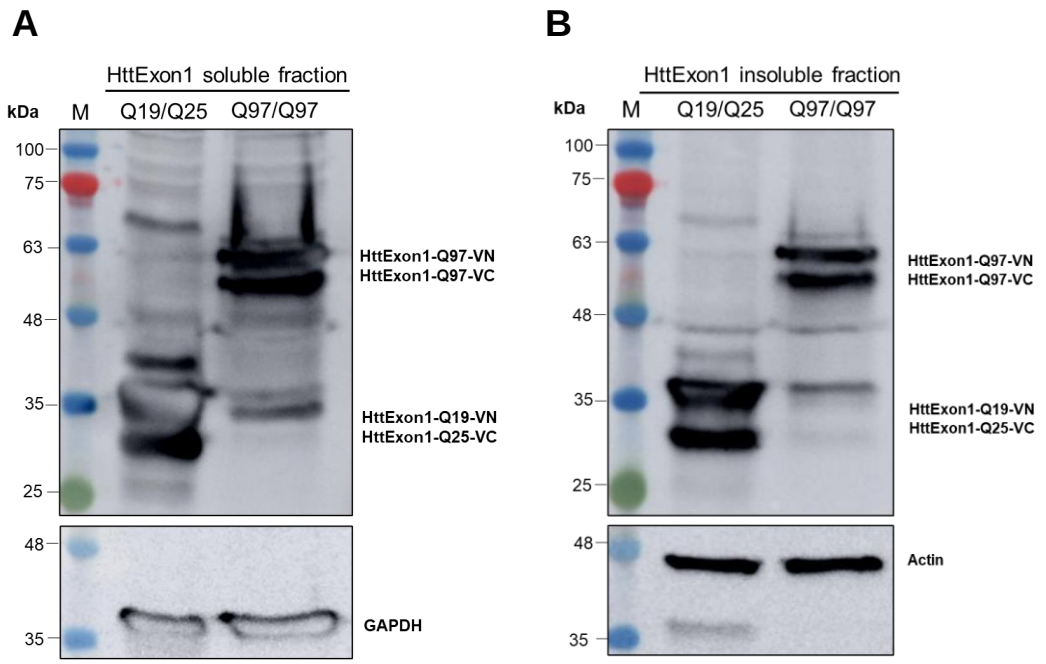

Figure 3.19 – Immunoblotting of wild-type and mutant HttExon1 after protein extraction from HeLa transfected cells. For protein extraction, 800.000 cells were seeded in 60 mm dish plates, transfected after 24h and maintained at 37ºC and 5%CO2 until the extraction was performed (24h after transfection). HeLa cells were transfected with 0.5 µg of HttExon1-19Q/25Q and 1 µg HttExon1-97Q/97Q plasmids. Both gels were loaded with 30 µg of protein extract after cell lysis. **(A)** The soluble fraction was obtained after extraction with native lysis buffer followed by centrifugation. Both wild-type and mutant constructs were expressed in cells after transfection. The analysis included the reference control GAPDH, which confirmed normal protein expression levels within the cell context. **(B)** The insoluble fraction was obtained after resuspension of the pellet obtained for the soluble fraction with denaturing lysis buffer, followed by centrifugation. The western blot confirmed the presence of both wild-type and mutant constructs in the insoluble fraction. The control used was actin, that revealed normal protein expression within the cell.

3.6.2. Investigation of the effect of S100B on HttExon1-Q97/Q97 aggregation in cells

The S100B protein is a calcium-binding alarmin, primarily synthetized by astrocytes within the brain. Recent discoveries have uncovered S100B's novel chaperone activity inhibiting amyloid aggregation and mitigating toxicity, particularly in the context of Alzheimer's Disease. Consequently, S100B has emerged as a novel player in maintaining brain proteostasis, extending its potential implications in other neurodegenerative diseases such as Huntington's Disease.

This investigation was driven by the following fundamental hypothesis: Since S100B inhibits amyloid aggregation in neurodegenerative-associated proteins, and with the observation that it also displays an effect on huntingtin aggregation *in vitro* in the FRASE assay, could this effect also manifest within live cells expressing huntingtin? To address this hypothesis, the BiFC technique was employed as previously described.

For the following tests, the focus shifted specifically to the mutant form, HttExon1-Q97/Q97, due to its pathological relevance. Three distinct experimental conditions were tested. First, HeLa cells were transfected only with the HttExon1-Q97/Q97 plasmids, serving as the control condition (Ø). Second, along with the mutant huntingtin constructs HttExon1-Q97/Q97, S100B fused with a fluorescent protein, cerulean, was also incorporated into the cells, denoted as S100B-Cerulean. S100B-Cerulean, in the context of this work, as well as the mutant huntingtin constructs, were considered as endogenous proteins, as the plasmids were introduced in the cells. Finally, the third condition tested HeLa cells transfected with HttExon1-Q97/Q97, following the addition of exogenous S100B to the medium –

specifically S100B-Myc protein. The Myc-tagged S100B protein enabled the differentiation between endogenous S100B expression and the externally introduced variant (S100B-Myc). S100B-Myc was added to the medium to a final concentration of 30 µM, a concentration demonstrated to preserve cell viability [90]. To facilitate the nomenclature, the three experiments will be referred to as control (Ø), Cerulean and Myc. The three conditions were tested following the same protocol.

Analysis of the microscope images revealed the presence of punctae aggregates and aggresomes within all three experimental setups. As an initial step, the total number of cells was counted, as well as the number of cells that contained any types of aggregates, whether punctae or aggresomes. The data is represented in a plot in **Figure 3.20**, in percentage of cells with aggregates normalized to the control condition. The percentage of cells with aggregates in Cerulean condition is 91,5%, compared to the control (100%). This result indicates that upon S100B-Cerulean transfection, the number of aggregates was reduced, suggesting that endogenous S100B might have a mitigating effect on aggregate formation, implying a potential role of S100B in regulating huntingtin aggregation dynamics. On the other hand, when comparing Myc with the control, the percentage of the number of cells with aggregates increased by 11%, with 110,9% of cells exhibiting any form of aggregation. Statistical analysis was conducted using ordinary one-way ANOVA, which however did not reveal statistically significant differences. While the precise mechanism behind this observation necessitates further investigation, it could potentially reflect complex interactions between the addition of exogenous S100B and huntingtin aggregation pathways in live cells.

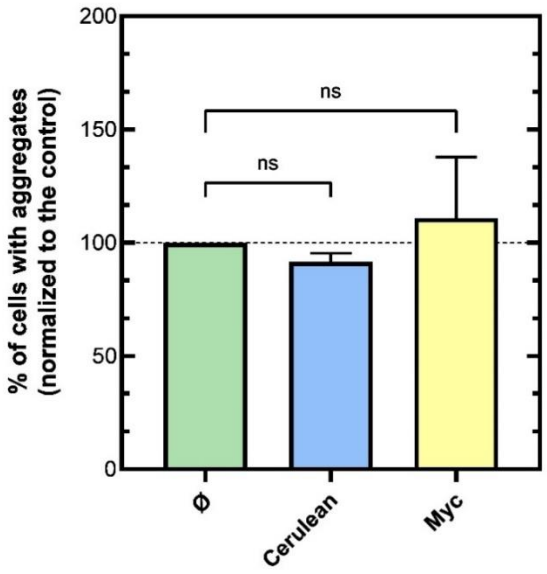

Figure 3.20 – Huntingtin aggregation patterns in live cells: Impact of endogenous S100B-Cerulean and exogenous S100B-Myc. This plot represents data derived from the observations of live cells expressing huntingtin under different conditions, using widefield fluorescence microscopy. The images were obtained 24h after plasmid transfection and the addition of the exogenous S100B-Myc. The investigation specifically targeted the mutant form, HttExon1-Q97/Q97, due to its pathological relevance. Three distinctive experimental conditions were meticulously examined: 1) Control condition (Ø): HeLa cells were solely transfected with HttExon1-Q97/Q97 plasmids, serving as the control. 2) S100B-Cerulean condition: In addition to the mutant huntingtin constructs, S100B fused with the fluorescent protein cerulean was introduced into the cells (S100B-Cerulean). Both S100B-Cerulean and the mutant huntingtin constructs were endogenous proteins, as their respective plasmids were introduced into the cells. 3) Addition of S100B-Myc condition: HeLa cells transfected with HttExon1-Q97/Q97 were subjected to exogenous S100B addition, specifically S100B-Myc protein. The Myc-tagged S100B enabled differentiation between endogenous and exogenous S100B. S100B-Myc was added to the medium at a concentration of 30 µM, a level that has been demonstrated to sustain cell viability [90].

Next, we wanted to look further into the types of aggregates within each different experimental condition tested, to better understand which type of aggregates were more or less expressed in each setup. Cells expressing a large aggresomes in the cytoplasm were categorized as aggresomes-positives, while those displaying punctae aggregates were classified as punctae-positives, and the respective counts were recorded accordingly. In **Figure 3.21A** and **Figure 3.21B**, the percentage of cells expressing punctae aggregates or aggresomes is shown, respectively.

In **Figure 3.21A**, the percentage of cells with aggresomes was very similar in the three setups. For Cerulean, the average of percentage of cells exhibiting large aggresomes was 99,9%, whereas for Myc this value increased to 106,4%. These preliminary results show that in these sets of experiments any form of S100B, whether endogenous or exogenous, have a relevant impact on the appearance of aggresomes within the cells. In **Figure 3.21B**, that depicts the percentage of cells with punctae aggregates, the contrary is observed. Comparing to the control (100%), both the Cerulean and Myc conditions exhibit a reduction in the population of cells displaying this particular type of aggregate structure, with percentages of 77,2% and 85,9%, respectively. Even though the ordinary one-way ANOVA used did not find these differences statistically significant, these results suggest that S100B might exert an effect on the early stages of punctae aggregate formation. These early formations might subsequently serve as precursors to the more developed dynamics culminating in the emergence of larger aggresomes structures.

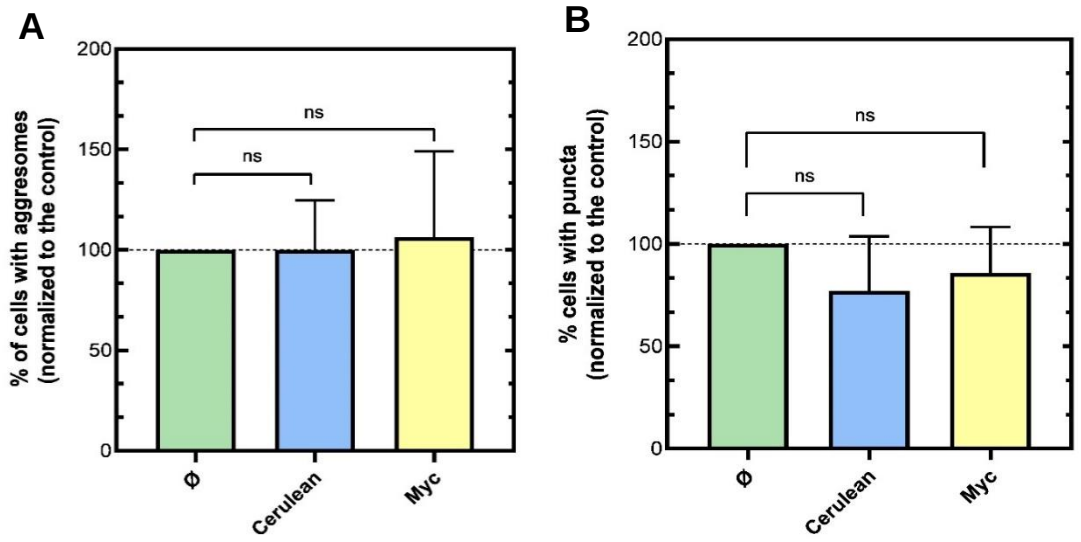

Figure 3.21 – Huntingtin aggregation patterns in live cells: Impact of endogenous S100B-Cerulean and exogenous S100B-Myc in aggresomes and punctae aggregates. (A) Percentage of cells exhibiting aggresomes, normalized to the control. The proportion of cells with aggresomes were strikingly similar across the three setups. Notably, for the Cerulean condition, an average of 99.9% of cells exhibited prominent aggresomes, while this percentage increased to 106.4% for the Myc condition. An ordinary one-way ANOVA was conducted, yielding no statistically significant differences between the control and the other two conditions tested**. (B) Percentage of cells exhibiting punctae aggregates, normalized to the control.** When contrasted with the control (100%), both the Myc and Cerulean conditions display a decrease in the population of cells presenting this specific aggregate morphology, with recorded percentages of 77.2% and 85.9%, respectively. Despite the absence of statistical significance in the ordinary one-way ANOVA analysis, these outcomes suggest a potential impact of S100B on the initial phases of punctae aggregate formation.

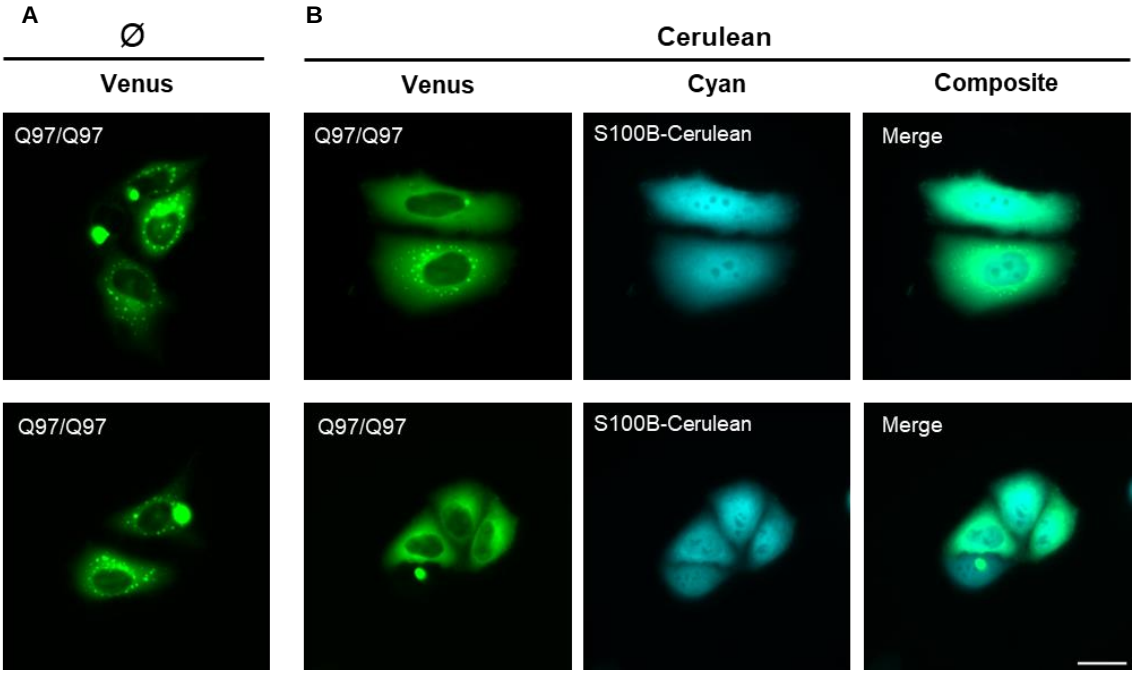

Figure 3.22 – Live-cell imaging of HeLa cells transfected with HttExon1-Q97/Q97 and with S100B-Cerulean plasmids in live HeLa cells using widefield fluorescence microscopy. Cells were seeded on ibiTreat dishes and co-transfected with BiFC pair constructs for 48h. The Venus channel captures the fluorescence emitted by the Venus-tagged HttExon1-Q97/Q97 fragments, revealing their distribution and interactions within the cell. The Cyan channel, highlights the expression of S100B-Cerulean, providing insights of its distribution in the cell. **(A)** The expression of HttExon1-Q97/Q97 on the control condition exhibit the appearance of both punctae aggregates and aggresomes. **(B)** The expression of HttExon1-Q97/Q97 with S100B-Cerulean decrease punctae aggregates formation comparing to the control (Ø). Scale bar = 30 µm.

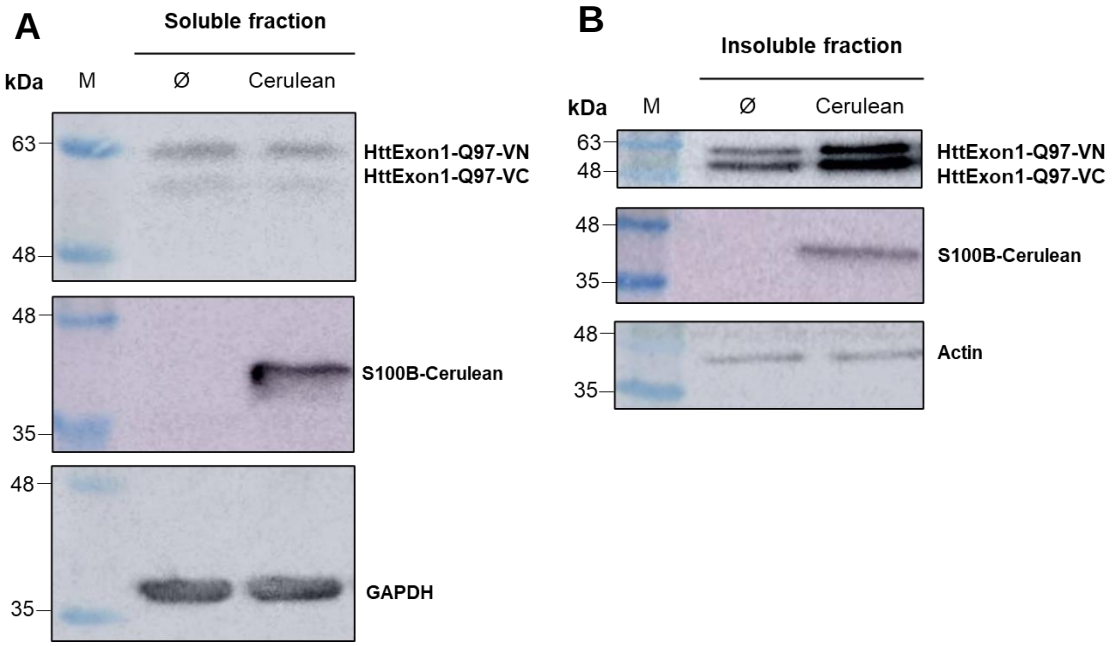

Figure 3.23 – Immunoblotting of wild-type and mutant HttExon1-Q97/Q97 after protein extraction from HeLa transfected cells. For protein extraction, 800.000 cells were seeded in 60 mm dish plates, transfected after 24h and maintained at 37ºC and 5%CO2 until the extraction was performed (24h after transfection). HeLa cells were transfected with 1 µg HttExon1-97Q/97Q and 1 µg of S100B-Cerulean plasmids. Both gels were loaded with 20 µg of protein extract after cell lysis. **(A)** The soluble fraction was obtained after extraction with native lysis buffer followed by centrifugation. Both the mutant and the S100B-Cerulean constructs were expressed in cells after transfection. The analysis included the reference control GAPDH, which confirmed normal protein expression levels within the cell context. **(B)** The insoluble fraction was obtained after resuspension of the pellet obtained for the soluble fraction with denaturing lysis buffer, followed by centrifugation. The western blot confirmed the presence of mutant and S100B-Cerulean constructs in the insoluble fraction. The control used was actin, that revealed normal protein expression within the cell.

To obtain the soluble fraction, the proteins were extracted by scrapping the Petri dishes using native lysis buffer, sonication, and a subsequent centrifugation step. Both the mutant construct and the S100B-Cerulean construct were expressed within the cellular soluble fraction subsequent to transfection, as observed in **Figure 3.23A**. GAPDH expression substantiated the normal levels of protein expression within the cellular environment. The insoluble fraction was isolated by resuspending the pellet derived from the soluble fraction in denaturing lysis buffer, followed by a second step of sonication, and succeeded by a centrifugation step. Also, through the western blot analysis, the presence of both the mutant and S100B-Cerulean constructs within the insoluble fraction was confirmed (**Figure 3.23B**). The control utilized in this investigation was actin, which the appearance of the specific band validated the usual protein expression within the cellular system, as described before.

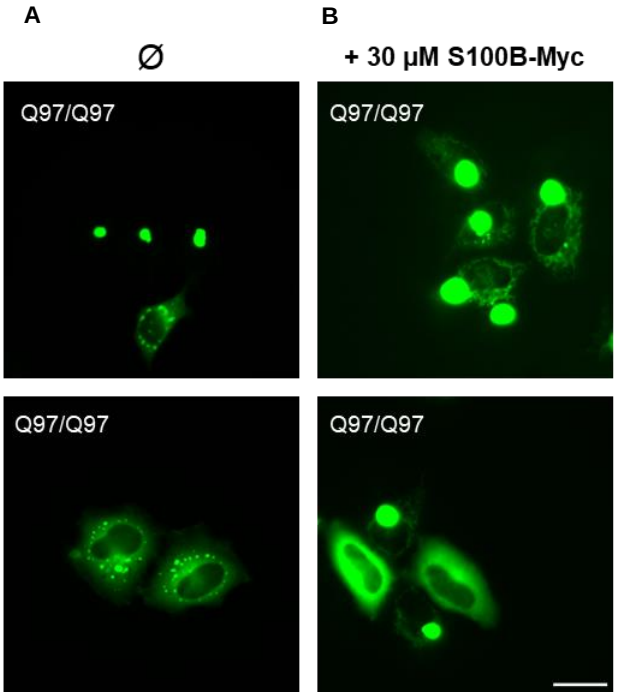

Figure 3.24 – Live-cell imaging of HeLa cells transfected with HttExon1-Q97/Q97 and upon S100B-Myc addition. Cells were seeded on ibiTreat dishes, co-transfected with BiFC pair constructs and S100B-Myc were added. Cells were incubated for 24h until microscope analysis. The Venus channel captures the fluorescence emitted by the Venus-tagged HttExon1-Q97/Q97 fragments, revealing their distribution and interactions within the cell in the two set of experiments shown. **(A)** The expression of HttExon1-Q97/Q97 on the control condition exhibit the appearance of both punctae aggregates and aggresomes. **(B)** The expression of HttExon1-Q97/Q97 with addition of 30 µM of S100B-Myc in the medium decrease punctae aggregates formation but show increasing appearance of aggresomes within the cytoplasmatic region comparing to the control (Ø). Scale bar = 30 µm.

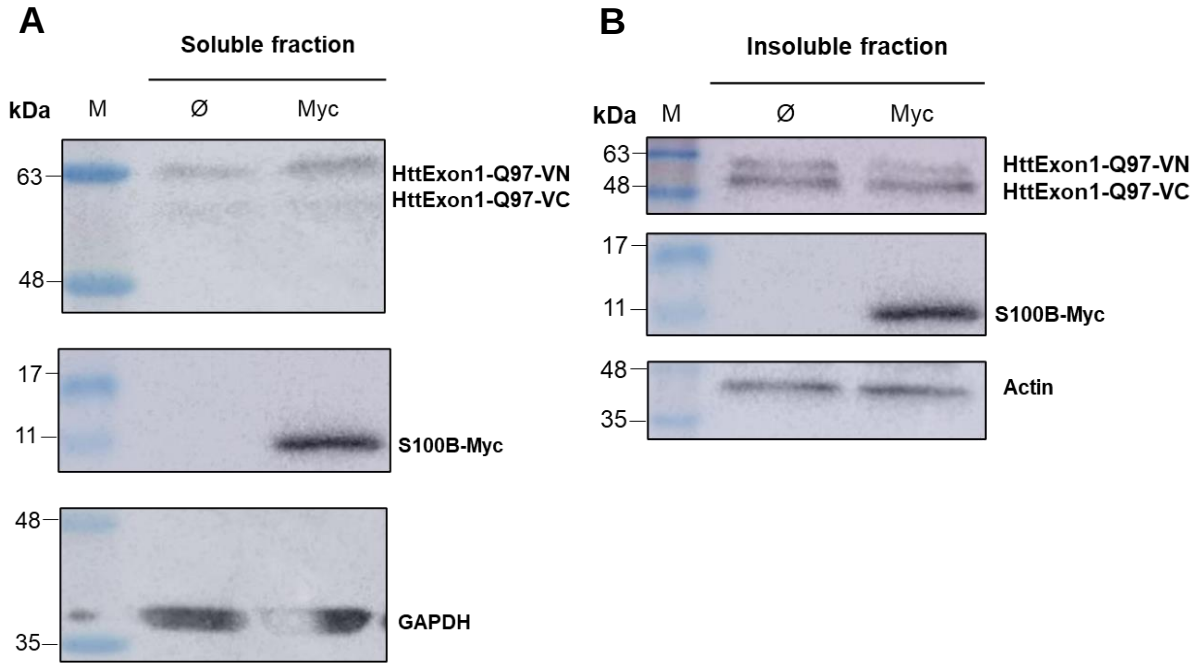

Figure 3.25 – Immunoblotting of wild-type and mutant HttExon1-Q97/Q97 after protein extraction from HeLa transfected cells upon S100B-Myc addition. The soluble and insoluble fraction were obtained as described before. **(A)** Immunoblotting of the soluble fraction confirmed the expression of HttExon1-Q97/Q97, as well as S100B-Myc uptake by cells. GAPDH detection confirmed normal protein expression within the cells. **(B)** Immunoblotting of the insoluble fraction also confirmed HttExon1-Q97/Q97 as well as incorporation of S100B-Myc in the cells. The control actin also showed normal protein expression.

For the Myc condition, the cells were transfected with HttExon1-Q97/Q97 as described before. After the transfection, S100B-Myc was added at a 30µM concentration, that preserved cell viability. S100B-Myc was introduced into the cellular medium, under the premise that the cells would uptake and internalize the protein, which was also addressed in the western blot analysis. Immunoblotting results showed that both the soluble and insoluble fractions of the cell lysates expressed both the HttExon1-Q97/Q97 constructs, and in addition contained the added S100B-Myc (**Figure 3.25**). This collective observation suggests the potential for all these proteins to be tethered to organelles, cellular membranes, or to manifest as aggregates, with a specific emphasis on the expressed HttExon1-Q97/Q97 fragments, but also to the S100B-Myc whose band appear in the insoluble fraction.

To deepen the investigation on how S100B affects huntingtin aggregation in a cellular context, the number of punctae per cell expressing punctae aggregates was quantified within each experimental condition, and plotted as it follows on **Figure 3.26**.

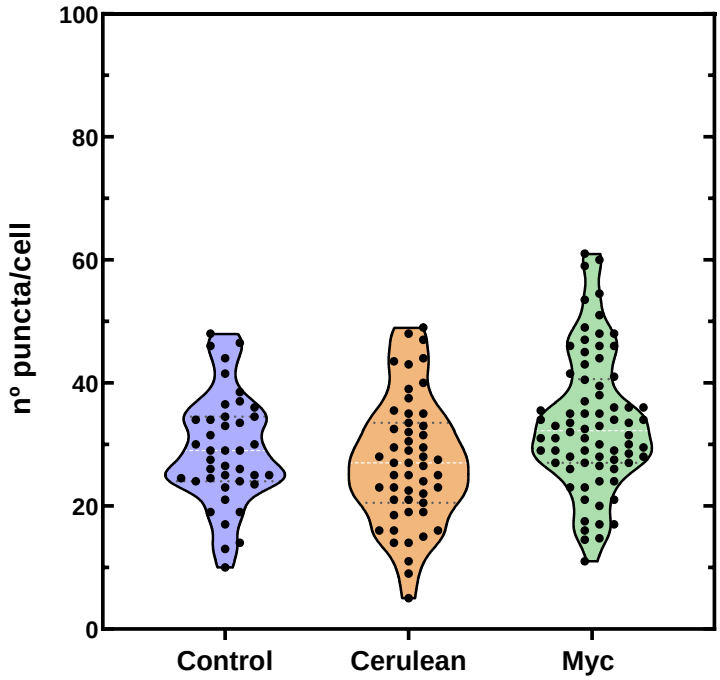

Figure 3.26 – Violin plots of the number of punctae per cell for each tested condition. The number of punctae per cell was meticulously counted across each experimental scenario, providing a quantitative measure of the extent of huntingtin aggregation in response to different conditions. The Myc condition exhibits higher values of quantified punctae per cell, while the Cerulean exhibits a similar distribution to the control.

Notably, a distinct trend in the distribution of quantified punctae per cell within the Myc and Cerulean conditions was observed, as compared to the control. The violin plot for the Myc condition, characterized by its elevated tail, underscores a pronounced increase in the puncta/cell count comparing to the control and to the Cerulean conditions. Conversely, the Cerulean condition reveals a shape and distribution that are similar to the control, suggesting that the presence of S100B-Cerulean does not significantly alter the punctae per cell count.
However, further experiments are required to clarify these preliminary results.

IV. Conclusion

Huntington's disease (HD) is a progressive neurodegenerative condition characterized by an autosomal dominant inheritance pattern, resulting in a combination of motor, cognitive, and behavioral symptoms. This incurable disease arises from an expansion of CAG trinucleotide repeats in the *HTT* gene, leading to an elongated polyglutamine tract within the huntingtin protein. This alteration gives rise to mutant huntingtin, causing destabilization in Htt's structure and promoting the formation of amyloidogenic protein aggregates, primarily composed of N-terminal Htt fragments. This accumulation disrupts crucial brain functions, ultimately leading to neuronal degeneration. S100B is a calcium-binding alarmin primarily produced by astrocytes within the nervous system. Depending on its concentration, S100B can function as an exacerbating pro-inflammatory cytokine or as a protective anti-aggregation chaperone. In the presence of calcium, it inhibits amyloid aggregation and mitigates toxicity. Recognizing S100B's significance in other neurodegenerative disorders and its role as a chaperone, this research aimed to investigate its potential involvement in huntingtin aggregation and its impact on the pathology of Huntington's disease.

For that, a Myc tag version of the S100B protein was produced to be used in this study. The further biophysical investigation and characterization of S100-Myc protein approach used integrated both circular dichroism spectroscopy and 8-anilino-1-naphthalenesulfonic acid (ANS) assays. The CD spectroscopy analysis allowed to investigate the secondary structural characteristics of the protein, providing insights into its folding patterns and stability. Notably, the far-UV CD spectra of S100B-Myc exhibited a consistent prevalence of α-helical secondary structure, which remained even under high temperatures, suggesting a remarkable stability of the protein. Building upon these findings, the ANS assays explored the protein's hydrophobic interactions and dynamic responses. With the introduction of Ca^{2+} ions, a recognized ligand for S100B, the ANS assays depicted evident shifts in fluorescence emission spectra. A blue shift in fluorescence emission peak and also a progressive increase in fluorescence intensity, upon the addition of increasing calcium concentrations from 1 mM to 2 mM, suggested conformational changes in the protein upon calcium binding.

As a first approach to investigate the effect of S100B in huntingtin aggregation, a FRET-based mHtt aggregate seeding (FRASE) assay was performed. Initially, various concentrations of S100B, ranging from 0.5 µM to 8 µM, were tested, while maintaining a constant concentration of 2 µM for HttExon1-Q48 (CyPet and YPet fused), with no addition of calcium. Across all conditions where S100B was added at different concentrations, no noticeable difference in huntingtin aggregation was observed. Given the hypothesis that S100B's chaperone-like function is dependent on the presence of calcium, a second FRASE assay was conducted, in the presence of calcium. This finding uncovered a significant result – the impact of the S100B-Ca^{2+} complex on inhibiting the aggregation dynamics of huntingtin. This investigation underscored S100B's role as a potent modulator of Htt aggregation and in its role on potentially mitigating the formation of toxic huntingtin aggregates, the hallmark of HD.

The LuTHy (Luminescence-based Two-Hybrid) assay was then employed to investigate direct interactions between Htt and specific interactors, including FEZ1, S100A1, and S100B. This bioluminescence-based two-hybrid method allowed for the sensitive and specific detection of protein-protein interactions (PPIs) in mammalian cells. BRET (bioluminescence resonance energy transfer) measurements were used to assess potential interactions. The interactions were considered "true" when cBRET (corrected BRET) values were equal or exceeded 0.01. The results revealed that most data points, including the positive control FEZ1, did not surpass the predefined cBRET threshold, raising

concerns about the assay's reliability in this set of experiments, as LuTHy is a well-established and fully validated assay.. Therefore, the LuTHy assay provided preliminary insights of potential interactions between certain huntingtin constructs and S100B, but the results presented were inconclusive and suggest the need for assay optimization to enhance accuracy and reliability.

The BiFC assay provided real-time insights into aggregate formation in live cells, enabling the examination of wild-type and mutant HttExon1 fragments' aggregation patterns in the presence and absence of S100B. The results revealed that S100B, whether endogenous (Cerulean condition) or exogenous (Myc condition), might play a role in influencing aggregation dynamics. In the Cerulean condition, there was a reduction in the percentage of cells with aggregates, with only 91.5% of cells showing aggregates compared to the control (100%). This indicated that endogenous S100B potentially mitigated aggregate formation. Conversely, the Myc condition showed a slight increase, with approximately 110.9% of cells exhibiting various forms of aggregation. Further analysis delved into the types of aggregates present in each condition. While aggresomes appeared consistently across all setups (approximately 99.9% for Cerulean and 106.4% for Myc), punctae aggregates showed a decrease in both the Cerulean (77.2%) and Myc (85.9%) conditions, hinting at a potential effect of S100B on early punctae aggregate formation. However, the results obtained for this assay were not statistically significant. To validate protein expression, soluble and insoluble protein fractions were extracted and analyzed. Western blot results confirmed the presence of HttExon1-Q97/Q97 constructs with S100B-Cerulean and S100B-Myc in both fractions, suggesting their potential association with cellular structures or aggregates in the case of Htt.

Collectively, these assays provide conclusive insights into the multifaceted role of S100B in huntingtin aggregation. While S100B may exert a mitigating effect on aggregate formation under specific conditions, its influence varies depending on the experimental context, including the presence of calcium ions. These findings underscore the complexity of the interplay between S100B and huntingtin aggregation and highlight S100B's potential as a regulatory factor in mitigating the formation of toxic huntingtin aggregates, a critical aspect of HD pathology. Further investigations are warranted to unravel the precise mechanisms and therapeutic potential of S100B in HD.

V. References

1. Sousa, A.M.M., et al., *Evolution of the Human Nervous System Function, Structure, and Development.* Cell, 2017. **170**(2): p. 226-247.
2. Azevedo, F.A., et al., *Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain.* J Comp Neurol, 2009. **513**(5): p. 532-41.
3. Ludwig, P.E., V. Reddy, and M. Varacallo, *Neuroanatomy, Central Nervous System (CNS)*, in *StatPearls*. 2023: Treasure Island (FL).
4. Aranda-Anzaldo, A., *The post-mitotic state in neurons correlates with a stable nuclear higher-order structure.* Commun Integr Biol, 2012. **5**(2): p. 134-9.
5. Huntington, G., *On chorea. George Huntington, M.D.* J Neuropsychiatry Clin Neurosci, 2003. **15**(1): p. 109-12.
6. Owecki, M.K. and A. Magowska, *George Huntington (1850-1916).* J Neurol, 2019. **266**(3): p. 793-795.
7. Adam, O.R. and J. Jankovic, *Symptomatic treatment of Huntington disease.* Neurotherapeutics, 2008. **5**(2): p. 181-97.
8. Ross, C.A., et al., *Huntington disease: natural history, biomarkers and prospects for therapeutics.* Nat Rev Neurol, 2014. **10**(4): p. 204-16.
9. Benatar, M., et al., *Preventing amyotrophic lateral sclerosis: insights from pre-symptomatic neurodegenerative diseases.* Brain, 2022. **145**(1): p. 27-44.
10. Killoran, A., et al., *Characterization of the Huntington intermediate CAG repeat expansion phenotype in PHAROS.* Neurology, 2013. **80**(22): p. 2022-7.
11. *Unified Huntington's Disease Rating Scale: reliability and consistency. Huntington Study Group.* Mov Disord, 1996. **11**(2): p. 136-42.
12. Ross, C.A., et al., *Movement Disorder Society Task Force Viewpoint: Huntington's Disease Diagnostic Categories.* Mov Disord Clin Pract, 2019. **6**(7): p. 541-546.
13. Shoulson, I. and S. Fahn, *Huntington disease: clinical care and evaluation.* Neurology, 1979. **29**(1): p. 1-3.
14. Gusella, J.F., et al., *A polymorphic DNA marker genetically linked to Huntington's disease.* Nature, 1983. **306**(5940): p. 234-8.
15. *A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. The Huntington's Disease Collaborative Research Group.* Cell, 1993. **72**(6): p. 971-83.
16. Telenius, H., et al., *Somatic and gonadal mosaicism of the Huntington disease gene CAG repeat in brain and sperm.* Nat Genet, 1994. **6**(4): p. 409-14.
17. Migliore, S., J. Jankovic, and F. Squitieri, *Genetic Counseling in Huntington's Disease: Potential New Challenges on Horizon?* Front Neurol, 2019. **10**: p. 453.
18. Wheeler, V.C., et al., *Factors associated with HD CAG repeat instability in Huntington disease.* J Med Genet, 2007. **44**(11): p. 695-701.
19. Gusella, J.F., M.E. MacDonald, and J.M. Lee, *Genetic modifiers of Huntington's disease.* Mov Disord, 2014. **29**(11): p. 1359-65.
20. Quigley, J., *Juvenile Huntington's Disease: Diagnostic and Treatment Considerations for the Psychiatrist.* Curr Psychiatry Rep, 2017. **19**(2): p. 9.
21. Gusella, J.F. and M.E. MacDonald, *Huntington's disease: the case for genetic modifiers.* Genome Med, 2009. **1**(8): p. 80.
22. Novati, A., H.P. Nguyen, and J. Schulze-Hentrich, *Environmental stimulation in Huntington disease patients and animal models.* Neurobiol Dis, 2022. **171**: p. 105725.
23. Medina, A., et al., *Prevalence and Incidence of Huntington's Disease: An Updated Systematic Review and Meta-Analysis.* Mov Disord, 2022. **37**(12): p. 2327-2335.
24. Pringsheim, T., et al., *The incidence and prevalence of Huntington's disease: a systematic review and meta-analysis.* Mov Disord, 2012. **27**(9): p. 1083-91.
25. Cattaneo, E., C. Zuccato, and M. Tartari, *Normal huntingtin function: an alternative approach to Huntington's disease.* Nat Rev Neurosci, 2005. **6**(12): p. 919-30.
26. Clayton Wiley, M.-F.C., et al., *Neurobiology of Brain Disorders (Second Edition)*. 2022, Academic Press,: S.l. p. 1 online resource.

27. Gutekunst, C.A., et al., *Identification and localization of huntingtin in brain and human lymphoblastoid cell lines with anti-fusion protein antibodies.* Proc Natl Acad Sci U S A, 1995. **92**(19): p. 8710-4.
28. Ehrnhoefer, D.E., L. Sutton, and M.R. Hayden, *Small changes, big impact: posttranslational modifications and function of huntingtin in Huntington disease.* Neuroscientist, 2011. **17**(5): p. 475-92.
29. Cariulo, C., et al., *Phosphorylation of huntingtin at residue T3 is decreased in Huntington's disease and modulates mutant huntingtin protein conformation.* Proc Natl Acad Sci U S A, 2017. **114**(50): p. E10809-E10818.
30. Watkin, E.E., et al., *Phosphorylation of mutant huntingtin at serine 116 modulates neuronal toxicity.* PLoS One, 2014. **9**(2): p. e88284.
31. Jarosinska, O.D. and S.G.D. Rudiger, *Molecular Strategies to Target Protein Aggregation in Huntington's Disease.* Front Mol Biosci, 2021. **8**: p. 769184.
32. Xu, D., et al., *Sequence and structural analyses of nuclear export signals in the NESdb database.* Mol Biol Cell, 2012. **23**(18): p. 3677-93.
33. Xia, J., et al., *Huntingtin contains a highly conserved nuclear export signal.* Hum Mol Genet, 2003. **12**(12): p. 1393-403.
34. Guo, Q., et al., *The cryo-electron microscopy structure of huntingtin.* Nature, 2018. **555**(7694): p. 117-120.
35. Harding, R.J., et al., *Huntingtin structure is orchestrated by HAP40 and shows a polyglutamine expansion-specific interaction with exon 1.* Commun Biol, 2021. **4**(1): p. 1374.
36. Nasir, J., et al., *Targeted disruption of the Huntington's disease gene results in embryonic lethality and behavioral and morphological changes in heterozygotes.* Cell, 1995. **81**(5): p. 811-23.
37. Wang, G., et al., *Ablation of huntingtin in adult neurons is nondeleterious but its depletion in young mice causes acute pancreatitis.* Proc Natl Acad Sci U S A, 2016. **113**(12): p. 3359-64.
38. Schulte, J. and J.T. Littleton, *The biological function of the Huntingtin protein and its relevance to Huntington's Disease pathology.* Curr Trends Neurol, 2011. **5**: p. 65-78.
39. Velier, J., et al., *Wild-type and mutant huntingtins function in vesicle trafficking in the secretory and endocytic pathways.* Exp Neurol, 1998. **152**(1): p. 34-40.
40. Waelter, S., et al., *The huntingtin interacting protein HIP1 is a clathrin and alpha-adaptin-binding protein involved in receptor-mediated endocytosis.* Hum Mol Genet, 2001. **10**(17): p. 1807-17.
41. Hoffner, G., P. Kahlem, and P. Djian, *Perinuclear localization of huntingtin as a consequence of its binding to microtubules through an interaction with beta-tubulin: relevance to Huntington's disease.* J Cell Sci, 2002. **115**(Pt 5): p. 941-8.
42. Zuccato, C., et al., *Loss of huntingtin-mediated BDNF gene transcription in Huntington's disease.* Science, 2001. **293**(5529): p. 493-8.
43. Podvin, S., et al., *Mutant Huntingtin Protein Interaction Map Implicates Dysregulation of Multiple Cellular Pathways in Neurodegeneration of Huntington's Disease.* J Huntingtons Dis, 2022. **11**(3): p. 243-267.
44. Mangiarini, L., et al., *Exon 1 of the HD gene with an expanded CAG repeat is sufficient to cause a progressive neurological phenotype in transgenic mice.* Cell, 1996. **87**(3): p. 493-506.
45. Barbaro, B.A., et al., *Comparative study of naturally occurring huntingtin fragments in Drosophila points to exon 1 as the most pathogenic species in Huntington's disease.* Hum Mol Genet, 2015. **24**(4): p. 913-25.
46. Aguzzi, A. and T. O'Connor, *Protein aggregation diseases: pathogenicity and therapeutic perspectives.* Nat Rev Drug Discov, 2010. **9**(3): p. 237-48.
47. Pallares, I. and S. Ventura, *The Transcription Terminator Rho: A First Bacterial Prion.* Trends Microbiol, 2017. **25**(6): p. 434-437.
48. Weids, A.J., et al., *Distinct stress conditions result in aggregation of proteins with similar properties.* Sci Rep, 2016. **6**: p. 24554.
49. Scherzinger, E., et al., *Self-assembly of polyglutamine-containing huntingtin fragments into amyloid-like fibrils: implications for Huntington's disease pathology.* Proc Natl Acad Sci U S A, 1999. **96**(8): p. 4604-9.

50. Lin, H.K., et al., *Fibril polymorphism affects immobilized non-amyloid flanking domains of huntingtin exon1 rather than its polyglutamine core.* Nat Commun, 2017. **8**: p. 15462.
51. Boatz, J.C., et al., *Protofilament Structure and Supramolecular Polymorphism of Aggregated Mutant Huntingtin Exon 1*. J Mol Biol, 2020. **432**(16): p. 4722-4744.
52. Wagner, A.S., et al., *Self-assembly of Mutant Huntingtin Exon-1 Fragments into Large Complex Fibrillar Structures Involves Nucleated Branching.* J Mol Biol, 2018. **430**(12): p. 1725-1744.
53. Chen, S., et al., *Polyglutamine aggregation behavior in vitro supports a recruitment mechanism of cytotoxicity.* J Mol Biol, 2001. **311**(1): p. 173-82.
54. Thakur, A.K., et al., *Polyglutamine disruption of the huntingtin exon 1 N terminus triggers a complex aggregation mechanism.* Nat Struct Mol Biol, 2009. **16**(4): p. 380-9.
55. Slepko, N., et al., *Normal-repeat-length polyglutamine peptides accelerate aggregation nucleation and cytotoxicity of expanded polyglutamine proteins.* Proc Natl Acad Sci U S A, 2006. **103**(39): p. 14367-72.
56. Peskett, T.R., et al., *A Liquid to Solid Phase Transition Underlying Pathological Huntingtin Exon1 Aggregation.* Mol Cell, 2018. **70**(4): p. 588-601 e6.
57. Ryu, H., et al., *The therapeutic role of creatine in Huntington's disease.* Pharmacol Ther, 2005. **108**(2): p. 193-207.
58. Jimenez-Sanchez, M., et al., *Huntington's Disease: Mechanisms of Pathogenesis and Therapeutic Strategies.* Cold Spring Harb Perspect Med, 2017. **7**(7).
59. Moore, B.W., *A soluble protein characteristic of the nervous system.* Biochem Biophys Res Commun, 1965. **19**(6): p. 739-44.
60. Sedaghat, F. and A. Notopoulos, *S100 protein family and its application in clinical practice.* Hippokratia, 2008. **12**(4): p. 198-204.
61. Fritz, G. and C.W. Heizmann, *3D Structures of the Calcium and Zinc Binding S100 Proteins*, in *Handbook of Metalloproteins*. 2004.
62. Gonzalez, L.L., K. Garrie, and M.D. Turner, *Role of S100 proteins in health and disease.* Biochimica et Biophysica Acta (BBA) - Molecular Cell Research, 2020. **1867**(6): p. 118677.
63. Osterloh, D., V.V. Ivanenkov, and V. Gerke, *Hydrophobic residues in the C-terminal region of S100A1 are essential for target protein binding butnot for dimerization.* Cell Calcium, 1998. **24**(2): p. 137-151.
64. Schwaller, B., *Cytosolic Ca(2+) Buffers Are Inherently Ca(2+) Signal Modulators.* Cold Spring Harb Perspect Biol, 2020. **12**(1).
65. Zimmer, D.B., et al., *The S100 protein family: history, function, and expression.* Brain Res Bull, 1995. **37**(4): p. 417-29.
66. Michetti, F., et al., *The S100B story: from biomarker to active factor in neural injury.* J Neurochem, 2019. **148**(2): p. 168-187.
67. Horvath, I., et al., *Co-aggregation of pro-inflammatory S100A9 with alpha-synuclein in Parkinson's disease: ex vivo and in vitro studies.* J Neuroinflammation, 2018. **15**(1): p. 172.
68. Yang, F., et al., *The Role of S100A6 in Human Diseases: Molecular Mechanisms and Therapeutic Potential.* Biomolecules, 2023. **13**(7).
69. Shapiro, L.A., L.A. Bialowas-McGoey, and P.M. Whitaker-Azmitia, *Effects of S100B on Serotonergic Plasticity and Neuroinflammation in the Hippocampus in Down Syndrome and Alzheimer's Disease: Studies in an S100B Overexpressing Mouse Model.* Cardiovasc Psychiatry Neurol, 2010. **2010**.
70. Sorci, G., et al., *S100B protein in tissue development, repair and regeneration.* World J Biol Chem, 2013. **4**(1): p. 1-12.
71. Prez, K.D. and L. Fan, *Structural Basis for S100B Interaction with its Target Proteins.* J Mol Genet Med, 2018. **12**(3).
72. Michetti, F., et al., *The S100B Protein: A Multifaceted Pathogenic Factor More Than a Biomarker.* Int J Mol Sci, 2023. **24**(11).
73. Donato, R., et al., *Functions of S100 proteins.* Curr Mol Med, 2013. **13**(1): p. 24-57.
74. Donato, R., et al., *S100B's double life: intracellular regulator and extracellular signal.* Biochim Biophys Acta, 2009. **1793**(6): p. 1008-22.
75. Faisal, M., et al., *Diagnostic performance of biomarker S100B and guideline adherence in routine care of mild head trauma.* Scand J Trauma Resusc Emerg Med, 2023. **31**(1): p. 3.

76. Nisbet, R.M. and J. Gotz, *Amyloid-beta and Tau in Alzheimer's Disease: Novel Pathomechanisms and Non-Pharmacological Treatment Strategies.* J Alzheimers Dis, 2018. **64**(s1): p. S517-S527.
77. Figueira, A.J., et al., *S100B chaperone multimers suppress the formation of oligomers during Abeta42 aggregation.* Front Neurosci, 2023. **17**: p. 1162741.
78. Cristovao, J.S. and C.M. Gomes, *S100 Proteins in Alzheimer's Disease.* Front Neurosci, 2019. **13**: p. 463.
79. Moreira, G.G., et al., *Dynamic interactions and Ca(2+)-binding modulate the holdase-type chaperone activity of S100B preventing tau aggregation and seeding.* Nat Commun, 2021. **12**(1): p. 6292.
80. Figueira, A.J., et al., *Tetramerization of the S100B Chaperone Spawns a Ca(2+) Independent Regulatory Surface that Enhances Anti-aggregation Activity and Client Specificity.* J Mol Biol, 2022. **434**(19): p. 167791.
81. Cristovao, J.S., et al., *The neuronal S100B protein is a calcium-tuned suppressor of amyloid-beta aggregation.* Sci Adv, 2018. **4**(6): p. eaaq1702.
82. Cristovao, J.S., et al., *Targeting S100B with Peptides Encoding Intrinsic Aggregation-Prone Sequence Segments.* Molecules, 2021. **26**(2).
83. Moreira, G.G. and C.M. Gomes, *Tau liquid-liquid phase separation is modulated by the Ca(2+) -switched chaperone activity of the S100B protein.* J Neurochem, 2023. **166**(1): p. 76-86.
84. Hulisz, D., *Amyotrophic lateral sclerosis: disease state overview.* Am J Manag Care, 2018. **24**(15 Suppl): p. S320-S326.
85. Serrano, A., et al., *The Astrocytic S100B Protein with Its Receptor RAGE Is Aberrantly Expressed in SOD1(G93A) Models, and Its Inhibition Decreases the Expression of Proinflammatory Genes.* Mediators Inflamm, 2017. **2017**: p. 1626204.
86. Kamma, E., et al., *Central nervous system macrophages in progressive multiple sclerosis: relationship to neurodegeneration and therapeutics.* J Neuroinflammation, 2022. **19**(1): p. 45.
87. Ragagnin, A.M.G., et al., *Motor Neuron Susceptibility in ALS/FTD.* Front Neurosci, 2019. **13**: p. 532.
88. Vieweg, S., et al., *An Intein-based Strategy for the Production of Tag-free Huntingtin Exon 1 Proteins Enables New Insights into the Polyglutamine Dependence of Httex1 Aggregation and Fibril Formation.* J Biol Chem, 2016. **291**(23): p. 12074-86.
89. Botelho, H.M., G. Fritz, and C.M. Gomes, *Analysis of S100 oligomers and amyloids.* Methods Mol Biol, 2012. **849**: p. 373-86.
90. Hagmeyer, S., et al., *Zinc Binding to S100B Affords Regulation of Trace Metal Homeostasis and Excitotoxicity in the Brain.* Front Mol Neurosci, 2017. **10**: p. 456.
91. Ast, A., et al., *mHTT Seeding Activity: A Marker of Disease Progression and Neurotoxicity in Models of Huntington's Disease.* Mol Cell, 2018. **71**(5): p. 675-688 e6.
92. Trepte, P., et al., *LuTHy: a double-readout bioluminescence-based two-hybrid technology for quantitative mapping of protein-protein interactions in mammalian cells.* Mol Syst Biol, 2018. **14**(7): p. e8071.
93. Branco-Santos, J., et al., *Protein phosphatase 1 regulates huntingtin exon 1 aggregation and toxicity.* Hum Mol Genet, 2017. **26**(19): p. 3763-3775.
94. Atwal, R.S., et al., *Kinase inhibitors modulate huntingtin cell localization and toxicity.* Nat Chem Biol, 2011. **7**(7): p. 453-60.
95. Shah, N.H. and T.W. Muir, *Inteins: Nature's Gift to Protein Chemists.* Chem Sci, 2014. **5**(1): p. 446-461.
96. Rath, A., et al., *Detergent binding explains anomalous SDS-PAGE migration of membrane proteins.* Proc Natl Acad Sci U S A, 2009. **106**(6): p. 1760-5.
97. Haase-Kohn, C., et al., *Copper-mediated cross-linking of S100A4, but not of S100A2, results in proinflammatory effects in melanoma cells.* Biochem Biophys Res Commun, 2011. **413**(3): p. 494-8.
98. Kelly, S.M. and N.C. Price, *Circular Dichroism: Studies of Proteins*, in *Encyclopedia of Life Sciences*. 2009.
99. Price, N.C., *Circular Dichroism: Studies of Proteins*, in *Encyclopedia of Life Sciences*. 2001.

100. Schonbrunn, E., et al., *Structural basis for the interaction of the fluorescence probe 8-anilino-1-naphthalene sulfonate (ANS) with the antibiotic target MurA.* Proc Natl Acad Sci U S A, 2000. **97**(12): p. 6345-9.
101. Gasymov, O.K. and B.J. Glasgow, *ANS fluorescence: potential to augment the identification of the external binding sites of proteins.* Biochim Biophys Acta, 2007. **1774**(3): p. 403-11.
102. Popiel, H.A., et al., *The Aggregation Inhibitor Peptide QBP1 as a Therapeutic Molecule for the Polyglutamine Neurodegenerative Diseases.* J Amino Acids, 2011. **2011**: p. 265084.
103. Razar, R., et al., *The importance of fasciculation and elongation protein zeta-1 in neural circuit establishment and neurological disorders.* Neural Regen Res, 2022. **17**(6): p. 1165-1171.
104. Chua, J.J., R. Jahn, and D.R. Klopfenstein, *Managing intracellular transport.* Worm, 2013. **2**(1): p. e21564.
105. Kobayashi, H., et al., *Bioluminescence resonance energy transfer-based imaging of protein-protein interactions in living cells.* Nat Protoc, 2019. **14**(4): p. 1084-1107.
106. Lai, H.T. and C.M. Chiang, *Bimolecular Fluorescence Complementation (BiFC) Assay for Direct Visualization of Protein-Protein Interaction in vivo.* Bio Protoc, 2013. **3**(20).
107. Kerppola, T.K., *Visualization of molecular interactions by fluorescence complementation.* Nat Rev Mol Cell Biol, 2006. **7**(6): p. 449-56.
108. Swaih, A.M., et al., *Kynurenine 3-Monooxygenase Interacts with Huntingtin at the Outer Mitochondrial Membrane.* Biomedicines, 2022. **10**(9).
109. Herrera, F., et al., *Visualization of cell-to-cell transmission of mutant huntingtin oligomers.* PLoS Curr, 2011. **3**: p. RRN1210.
110. Sikand, K., et al., *Housekeeping gene selection advisory: glyceraldehyde-3-phosphate dehydrogenase (GAPDH) and beta-actin are targets of miR-644a.* PLoS One, 2012. **7**(10): p. e47510.